WITH COM
NEW DELHI.
THE ASIA FOUNDATION

CLASSICAL
ELECTRICITY and MAGNETISM

ADDISON-WESLEY SERIES IN ADVANCED PHYSICS

Feenberg and Pake — Quantum Theory of Angular Momentum
Goldstein — Classical Mechanics
Jauch and Rohrlich — The Theory of Photons and Electrons
Landau and Lifshitz — The Classical Theory of Fields
Panofsky and Phillips — Classical Electricity and Magnetism
Sachs — Nuclear Theory

CLASSICAL ELECTRICITY and MAGNETISM

---

by

WOLFGANG K. H. PANOFSKY

Department of Physics
Stanford University

and

MELBA PHILLIPS

ADDISON-WESLEY PUBLISHING COMPANY, Inc.

READING, MASSACHUSETTS

ABAD AGRICULTURAL INSTITU
LIBRARY

ADDISON-WESLEY PUBLISHING COMPANY, Inc.

Printed in the United States of America

ALL RIGHTS RESERVED. THIS BOOK, OR PARTS THEREOF, MAY NOT BE REPRODUCED IN ANY FORM WITHOUT WRITTEN PERMISSION OF THE PUBLISHERS.

Library of Congress Catalog Card No. 55-5029

Second printing—August, 1929

537.1
P149C

181
4.8.66

LIBRARY
INSTITUTE ● ALLAHABAD
GIFT

PREFACE

This book is designed to emphasize those aspects of classical electricity and magnetism most useful to the modern student as a background both for experimental physics and for the quantum theory of matter and radiation. We have made no attempts at novelty beyond those inherent in looking at subject matter that has become a part of the foundations of physics, and has thus gained in usefulness as it has lost in immediacy. While no rigid adherence to historical development is attempted, the emphasis is on physical theory as evolved from fundamental empirical laws rather than on mathematics and strict internal logic. Thus Maxwell's equations are derived from the experimental laws of Coulomb, Ampere, and Faraday, instead of being postulated initially. In the opinion of the authors the physical concepts emerge more clearly in this way, and the approach represents the manner in which physical theory evolves in practice. The field formulation is preferred to the action-at-a-distance viewpoint even in electrostatics, however, since for the conventional treatment it is more readily extended to the nonstatic case. This despite the fact that it is possible, both for static and for nonstatic phenomena, to formulate an entirely consistent electromagnetic theory based on the delayed-action-at-a-distance principle.

The climax of 19th century electrodynamics was the theory of electromagnetic waves and its confirmation, and it is inevitable that any treatment of the subject today includes the principles of recent applications involving metallic boundaries. The introduction of the electrodynamic potentials and the Hertz solution of the wave equation are treated in the conventional way, but we have chosen to introduce the special theory of relativity before undertaking the theory of the electron. Historically the evidence was building up simultaneously along two separate lines, and many of the early difficulties in the derivation of radiation theory as applied to elementary charges were clarified in a very simple way by relativistic considerations. This approach has the advantage that the other problems of classical electron theory, especially those which have taken on added significance with the advent of quantum theory, can be exhibited more clearly.

Rationalized mks units are used throughout, simply because the majority of modern reference books and papers are now written in this system. Especially in the consideration of the electron, all quantities are so written that they can be immediately translated into Gaussian units. In

Appendix I will be found a discussion of the units in current use, and tables contain the fundamental relations of electrodynamics expressed in various systems as well as numerical conversion factors.

The text is based on graduate course lectures given by one of us (Panofsky) at the University of California and Stanford University. Early mimeographed notes on much of the subject matter were prepared with the aid of Howard Chang, Roger Wallace, Richard Madey, and Lee Aamodt, whose help is gratefully acknowledged. The editorial help of Miss Laurose Becker is also acknowledged with thanks.

The reader is assumed to have had courses in advanced calculus, differential equations, vector analysis, and, at least for the latter portions, is assumed to be familiar with classical mechanics on the graduate level. Prior knowledge of tensor analysis would be helpful, but is not necessary. References to appropriate collateral and background material are included at the end of each chapter, with some indication of what relevant material is to be found in each reference, and a full bibliographical list is given at the end.

The presentation is designed to be somewhat flexible, depending on the organization of course material. For purely theoretical courses Chapters 4 and 5, together with portions of other chapters dealing with particular applications of potential theory, etc., may be omitted entirely. Some of the material in Chapter 11 is often covered in optics courses. And if a course in relativity theory is given separately Chapters 14–17 may be omitted, since we have endeavored to make Chapter 18 continuous with Chapter 13, insofar as the theory of radiation is concerned.

A final word about problems: for the most part they are designed to supplement the text. It had been our intention to give credit to original sources for those we did not invent ourselves, but in almost every case this turns out to be impossible: like discoveries, problems are rarely made singly, and in a subject as old as this ingenuity mainly recreates old ideas. And despite our adherence to the exhortation used by Becker, "be ye doers of the word and not hearers only, deceiving your own selves," we have not concentrated primarily on problem solving. The heart of the matter, we believe, lies in the ideas and their development.

CONTENTS

CHAPTER 1. THE ELECTROSTATIC FIELD IN VACUUM 1

1–1 Vector fields 1
1–2 The electric field 6
1–3 Coulomb's law 7
1–4 The electrostatic potential 9
1–5 The potential in terms of charge distribution 10
1–6 Field singularities 11
1–7 Clusters of point charges 12
1–8 Dipole interactions 15
1–9 Surface singularities 17
1–10 Volume distributions of dipole moment 20

CHAPTER 2. BOUNDARY CONDITIONS AND RELATION OF MICROSCOPIC TO MACROSCOPIC FIELDS 24

2–1 The displacement vector 24
2–2 Boundary conditions 26
2–3 The electric field in a material medium 29
2–4 Polarizability 33

CHAPTER 3. GENERAL METHODS FOR THE SOLUTION OF POTENTIAL PROBLEMS 37

3–1 Uniqueness theorem 37
3–2 Green's reciprocation theorem 38
3–3 Solution by Green's function 39
3–4 Solution by inversion 42
3–5 Solution by electrical images 44
3–6 Solution of Laplace's equation by the separation of variables . . . 45

CHAPTER 4. TWO-DIMENSIONAL POTENTIAL PROBLEMS 54

4–1 Conjugate complex functions 54
4–2 Capacity and field strength 56
4–3 The potential of a uniform field 57
4–4 The potential of a line charge 57
4–5 Complex transformations 58
4–6 General Schwarz transformation 59
4–7 Single-angle transformations 62
4–8 Multiple-angle transformations 64
4–9 Direct solution of Laplace's equation by the method of harmonics . 65
4–10 Illustration: Line charge and dielectric cylinder 66
4–11 Line charge in an angle between two conductors 69

CHAPTER 5. Three-dimensional Potential Problems 73
5–1 The solution of Laplace's equation in spherical coordinates 73
5–2 The potential of a point charge 74
5–3 The potential of a dielectric sphere and a point charge 75
5–4 The potential of a dielectric sphere in a uniform field 76
5–5 The potential of an arbitrary axially-symmetric spherical potential distribution 78
5–6 The potential of a charged ring 78
5–7 Problems not having axial symmetry 79
5–8 The solution of Laplace's equation in cylindrical coordinates 80
5–9 Application of cylindrical solutions to potential problems 82

CHAPTER 6. Energy Relations and Forces in the Electrostatic Field 86
6–1 Field energy in free space 86
6–2 Energy density within a dielectric 89
6–3 Thermodynamic interpretation of U 90
6–4 Thomson's theorem 91
6–5 Volume forces in the electrostatic field 92
6–6 Maxwell stress tensor 95
6–7 The behavior of dielectric liquids in an electrostatic field 100

CHAPTER 7. Currents and Their Interaction 107
7–1 Ohm's law 107
7–2 Electromotive force 108
7–3 The solution of stationary current problems 109
7–4 Time of relaxation in a homogeneous medium 111
7–5 The magnetic interaction of steady currents 112
7–6 The magnetic induction field 113
7–7 The magnetic scalar potential 114
7–8 The magnetic vector potential 116
7–9 Types of currents 117
7–10 Polarization currents 118
7–11 Magnetization currents 119
7–12 Vacuum displacement current 121

CHAPTER 8. Magnetic Materials and Boundary Value Problems 124
8–1 Magnetic field intensity 124
8–2 Magnetic sources 125
8–3 Permeable media: magnetic susceptibility and boundary conditions 129
8–4 Magnetic circuits 130
8–5 Solution of boundary value problems by magnetic scalar potentials 131
8–6 Uniqueness theorem for the vector potential 132
8–7 The use of the vector potential in the solution of problems 133
8–8 The vector potential in two dimensions 135
8–9 The vector potential in cylindrical coordinates 138

CHAPTER 9. Maxwell's Equations 142
9–1 Faraday's law of induction 142
9–2 Maxwell's equations for stationary media 143

9–3 Faraday's law for moving media 144
9–4 Maxwell's equations for moving media 147
9–5 Motion of a conductor in a magnetic field 148

CHAPTER 10. ENERGY, FORCE, AND MOMENTUM RELATIONS IN THE ELECTROMAGNETIC FIELD 152
10–1 Energy relations in quasi-stationary current systems 152
10–2 Forces on current systems 154
10–3 Inductance 156
10–4 Magnetic volume force 159
10–5 General expressions for electromagnetic energy 160
10–6 Momentum balance 162

CHAPTER 11. THE WAVE EQUATION AND PLANE WAVES 166
11–1 The wave equation 166
11–2 Plane waves 168
11–3 Radiation pressure 172
11–4 Plane waves in a moving medium 174
11–5 Reflection and refraction at a plane boundary 176
11–6 Waves in conducting media and metallic reflection 181
11–7 General solutions of the homogeneous wave equation 183

CHAPTER 12. WAVE SOLUTIONS IN THE PRESENCE OF METALLIC BOUNDARIES 185
12–1 The nature of metallic boundary conditions 185
12–2 Eigenfunctions and eigenvalues of the wave equation 187
12–3 Cavities with rectangular boundaries 190
12–4 Cylindrical cavities 192
12–5 Circular cylindrical cavities 195
12–6 Wave guides 196
12–7 Scattering by a circular cylinder 199
12–8 Spherical waves 201
12–9 Scattering by a sphere 203

CHAPTER 13. THE INHOMOGENEOUS WAVE EQUATION 210
13–1 The wave equation for the potentials 210
13–2 Solution by Fourier analysis 212
13–3 Fourier components of the fields and radiation 215
13–4 The Hertz potential 219
13–5 Computation of radiation fields by the Hertz method 220
13–6 Electric dipole radiation 222
13–7 Multipole radiation 225

CHAPTER 14. THE EXPERIMENTAL BASIS FOR THE THEORY OF SPECIAL RELATIVITY 230
14–1 Galilean relativity and electrodynamics 230
14–2 The search for an absolute ether frame 231
14–3 The Lorentz-Fitzgerald contraction hypothesis 236

14–4 "Ether drag" . . . 236
14–5 Emission theories . . . 238
14–6 Summary . . . 239

CHAPTER 15. RELATIVISTIC KINEMATICS AND THE LORENTZ TRANSFORMATION . . . 243
15–1 The velocity of light and simultaneity . . . 243
15–2 Kinematic relations in special relativity . . . 244
15–3 The Lorentz transformation . . . 250
15–4 Geometric interpretations of the Lorentz transformation . . . 254
15–5 Transformation equations for velocity . . . 257

CHAPTER 16. COVARIANCE AND RELATIVISTIC MECHANICS . . . 261
16–1 The Lorentz transformation of a four-vector . . . 261
16–2 Some tensor relations useful in special relativity . . . 263
16–3 The conservation of momentum . . . 267
16–4 Relation of energy to momentum and to mass . . . 269
16–5 The Minkowski force . . . 272
16–6 The collision of two similar particles . . . 273

CHAPTER 17. COVARIANT FORMULATION OF VACUUM ELECTRODYNAMICS . 277
17–1 The four-vector potential . . . 277
17–2 The electromagnetic field tensor . . . 280
17–3 The Lorentz force . . . 284

CHAPTER 18. THE LIÉNARD-WIECHERT POTENTIALS AND THE FIELD OF A UNIFORMLY MOVING ELECTRON . . . 286
18–1 The Liénard-Wiechert potentials . . . 286
18–2 The fields of a charge in uniform motion . . . 289
18–3 Direct solution of the wave equation . . . 291
18–4 The "convection potential" . . . 293
18–5 The virtual photon concept . . . 294

CHAPTER 19. RADIATION FROM AN ACCELERATED CHARGE . . . 297
19–1 Fields of an accelerated charge . . . 297
19–2 Radiation at low velocity . . . 301
19–3 The case of $\dot{u}$ parallel to u . . . 301
19–4 Radiation when the acceleration is perpendicular to the velocity . . 306
19–5 Radiation with no restrictions on the acceleration or velocity . . . 307
19–6 Classical cross section for bremsstrahlung in a Coulomb field . . . 308
19–7 Čerenkov radiation . . . 309

CHAPTER 20. RADIATION REACTION . . . 314
20–1 Electromagnetic mass . . . 314
20–2 The reaction necessary to conserve radiated energy . . . 317
20–3 Computation of radiation reaction from the fields . . . 318
20–4 Difficulties in classical electron theory . . . 320

CHAPTER 21. RADIATION, SCATTERING, AND DISPERSION 322
21–1 Radiative damping of a charged harmonic oscillator 322
21–2 Forced vibrations 324
21–3 Scattering by an individual free electron 325
21–4 Scattering by a bound electron 327
21–5 Effect of a volume distribution of scatterers 328
21–6 Velocities of propagation 330
21–7 Scattering from a volume distribution. Rayleigh scattering . . . 332
21–8 Absorption of radiation by an oscillator 333
21–9 Equilibrium between an oscillator and a radiation field 336

CHAPTER 22. COVARIANT FORMULATION OF THE FIELD EQUATIONS IN MATERIAL MEDIA AND THE CONSERVATION LAWS OF ELECTRODYNAMICS 339
22–1 Covariant description of sources 339
22–2 The field equations in a material medium 341
22–3 Transformation properties of the partial fields 342
22–4 Covariant formulation of the conservation laws of vacuum electrodynamics 345
22–5 Some consequences of covariant formulation of the conservation laws . 346
22–6 The electromagnetic energy momentum tensor in material media . . 349

CHAPTER 23. THE MOTION OF CHARGED PARTICLES IN ELECTROMAGNETIC FIELDS 351
23–1 World-line description 351
23–2 Hamiltonian formulation and the transition to three-dimensional formalism 353
23–3 Equations for the trajectories 355
23–4 Applications 358

CHAPTER 24. HAMILTONIAN FORMULATION OF MAXWELL'S EQUATIONS . 364
24–1 Transition to a one-dimensional continuous system 364
24–2 Generalization to a three-dimensional continuum 366
24–3 The electromagnetic field 369
24–4 Periodic solutions in a box. Plane wave representation 372

APPENDIX I. UNITS AND DIMENSIONS IN ELECTROMAGNETIC THEORY . . 375
Tables: I–1. Fundamental Electromagnetic Relations Valid *in vacuo* as They Appear in the Various Systems of Units 382
I–2. Conversion Factors 384
I–3. Definition of Fields from Sources (mks system) 385
I–4. Useful Numerical Relations 386

APPENDIX II. VECTOR RELATIONS IN CURVILINEAR COORDINATES . . . 387
Table II–1. Coordinate Systems 389

BIBLIOGRAPHY 391

INDEX 395

CHAPTER 1

THE ELECTROSTATIC FIELD IN VACUUM

The interaction between material bodies can be described either by formulating the action at a distance between the interacting bodies or by separating the interaction process into the production of a *field* by one system and the action of the field on another system. These two alternative descriptions are physically indistinguishable in the static case. If the bodies are in motion, however, and the velocity of propagation of the interaction is finite, it is both physically and mathematically advantageous to ascribe physical reality to the field itself, even though it is possible in the electromagnetic case to replace the field concept by that of "delayed direct interaction." Moreover, the concept of field as describing the physical condition of space has played a key role in the historical development of electromagnetic theory. We shall therefore formulate even the electrostatic interactions as a field theory, which can then be extended to the consideration of nonstatic cases.

1–1 Vector fields. Field theories applicable to various types of interaction differ by the number of parameters necessary to define the field and the symmetry character of the field. The electric field is a three-dimensional vector field, i.e., a field definable by the specification of three components. The term "vector field" implies that the field components transform like coordinate intervals under transformations of the coordinates. Electric fields are also linear fields, which obey the principle of superposition, i.e., two or more fields may be added geometrically to determine the resultant field. The theory of vector fields was developed in connection with the study of fluid motion, a fact which is betrayed repeatedly by the vocabulary of the theory. We shall consider some general mathematical properties of such fields before specifying the physical content of the vectors.

All vector fields are uniquely defined if their circulation densities and source densities are given functions of the coordinates at all points in space, and if the totality of sources, as well as the source density, is zero at infinity. Let us prove this theorem formally. Consider a three-dimensional vector field $\mathbf{V}(x,y,z)$ such that

$$\nabla \cdot \mathbf{V} = s, \tag{1–1}$$

$$\nabla \times \mathbf{V} = \mathbf{c}. \tag{1–2}$$

We shall first show that if

$$\mathbf{V} = -\nabla\phi + \nabla \times \mathbf{A}, \tag{1-3}$$

where

$$\phi(x_\alpha) = \frac{1}{4\pi}\int \frac{s(x'_\alpha)}{r(x_\alpha,x'_\alpha)}\,dv' \tag{1-4}$$

and

$$\mathbf{A}(x_\alpha) = \frac{1}{4\pi}\int \frac{\mathbf{c}(x'_\alpha)}{r(x_\alpha,x'_\alpha)}\,dv', \tag{1-5}$$

then $\mathbf{V}$ satisfies Eqs. (1-1) and (1-2).

It is necessary to examine the notation of Eqs. (1-4) and (1-5) before proceeding with the proof. The symbol x_α stands for x, y, z at the *field point;* the symbol x'_α stands for x', y', z' at the *source point;* the function $r(x_\alpha,x'_\alpha)$ is the symmetric function

$$r(x_\alpha,x'_\alpha) = \left|\sqrt{\sum_{\alpha=1}^{\alpha=3}(x_\alpha - x'_\alpha)^2}\right|$$

representing the positive distance between field and source point. The reader should note carefully the functional relationships explicit in Eqs. (1-4) and (1-5). In integrals of this type these functional dependences will often not be fully stated; for example, we may write the volume integrals

$$\phi = \frac{1}{4\pi}\int \frac{s}{r}\,dv' \tag{1-4$'$}$$

$$\mathbf{A} = \frac{1}{4\pi}\int \frac{\mathbf{c}}{r}\,dv', \tag{1-5$'$}$$

as a short notation.

Let us demonstrate that $\mathbf{V}$ as expressed by Eq. (1-3) is a solution of Eqs. (1-1) and (1-2):

$$\nabla\cdot\mathbf{V} = -\nabla^2\phi + \nabla\cdot(\nabla\times\mathbf{A}) = -\nabla^2\phi$$
$$= -\frac{1}{4\pi}\nabla^2\left\{\int\frac{s}{r}\,dv'\right\}.$$

The Laplacian operator ∇^2 operates on the field coordinates; hence

$$\nabla\cdot\mathbf{V} = -\frac{1}{4\pi}\int s\nabla^2\left(\frac{1}{r}\right)dv'. \tag{1-6}$$

Now we can show that

$$\nabla^2 \left\{ \frac{1}{r(x_\alpha, x'_\alpha)} \right\} = -4\pi\, \delta(r), \tag{1–7}$$

where $\delta(r)$, the Dirac δ-function, is defined by the functional properties

$$\delta(r) = 0, \quad r \neq 0, \tag{1–8}$$

$$\int \delta(r)\, dv' = 1, \tag{1–9}$$

if the point $r = 0$ is included in the volume of integration, and by

$$\int f(x'_\alpha)\, \delta(r)\, dv' = f(x_\alpha), \tag{1–10}$$

for any arbitrary function f so long as the volume of integration includes the point $r = 0$. The δ-function is not an analytic function but essentially a notation for the functional properties of the three defining equations. It will always be used in terms of these properties.

Since it is evident by direct differentiation that $\nabla^2(1/r) = 0$ for $r \neq 0$, we have only to prove that

$$\int \nabla^2(1/r)\, dv' = -4\pi \tag{1–11}$$

in order to verify Eq. (1–7). [In Eq. (1–11) the point $r = 0$, that is, $x_\alpha = x'_\alpha$, is included in the volume of integration.] By the application of Gauss's divergence theorem, applicable to any vector $\mathbf{V}$,

$$\int \nabla \cdot \mathbf{V}\, dv = \int \mathbf{V} \cdot d\mathbf{S},$$

it is seen that

$$\int \nabla^2 \left(\frac{1}{r}\right) dv' = \int \nabla \left(\frac{1}{r}\right) \cdot d\mathbf{S}'$$

$$= -\int \frac{\mathbf{r} \cdot d\mathbf{S}'}{r^3} = -\int d\Omega,$$

where Ω is the solid angle subtended at x_α by the surface of integration S' over the variables x'_α. Since S' includes x_α, we have simply $\int d\Omega = 4\pi$, and Eq. (1–11) is verified. Hence from Eqs. (1–6) and (1–10),

$$\nabla \cdot \mathbf{V} = -\frac{1}{4\pi} \int s \nabla^2 \left(\frac{1}{r}\right) dv' = \int s(x'_\alpha)\, \delta(r)\, dv' = s(x_\alpha), \tag{1–12}$$

which was to be proved.

Similarly,

$$\nabla \times \mathbf{V} = -\nabla \times \nabla\phi + \nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2\mathbf{A}$$

$$= \frac{1}{4\pi}\left\{\int (\mathbf{c}\cdot\nabla)\nabla\left(\frac{1}{r}\right) dv' - \int \mathbf{c}\nabla^2\left(\frac{1}{r}\right) dv'\right\}. \tag{1-13}$$

We shall be able to show that the first integral vanishes if $\mathbf{c}$ is bounded in space. If we anticipate this result, we see immediately, from Eq. (1-7), that

$$\nabla \times \mathbf{V} = \int \mathbf{c}(x'_\alpha)\, \delta(r)\, dv' = \mathbf{c}(x_\alpha), \tag{1-14}$$

so that Eq. (1-2) is also satisfied.

To prove that the first term of Eq. (1-13) vanishes, let us examine the coordinate variables involved in the integrand. The operator ∇ has the components $\partial/\partial x_\alpha$. If we introduce the operator $\nabla'_\alpha = \partial/\partial x'_\alpha$, operating on the source coordinates, then for any arbitrary function $g[r(x_\alpha, x'_\alpha)]$, we have

$$\nabla g = -\nabla' g. \tag{1-15}$$

Therefore the first integral of Eq. (1-13) may be written

$$\mathbf{I} = \int (\mathbf{c}\cdot\nabla)\nabla\left(\frac{1}{r}\right) dv' = \int (\mathbf{c}\cdot\nabla')\nabla'\left(\frac{1}{r}\right) dv'.$$

The differential operators now operate on the variables of integration and we may integrate by parts. Each component of $\mathbf{I}$ becomes

$$I_\alpha = \int (\mathbf{c}\cdot\nabla')\frac{\partial}{\partial x'_\alpha}\left(\frac{1}{r}\right) dv'$$

$$= \int \nabla' \cdot \left\{\mathbf{c}\frac{\partial}{\partial x'_\alpha}\left(\frac{1}{r}\right)\right\} dv' - \int (\nabla'\cdot\mathbf{c})\frac{\partial}{\partial x'_\alpha}\left(\frac{1}{r}\right) dv'. \tag{1-16}$$

The second integral vanishes because the divergence of $\mathbf{c}$ is zero, since $\mathbf{c} = \nabla \times \mathbf{V}$. The first term can be transformed to a surface integral by means of Gauss's theorem; if $\mathbf{c}$ is bounded in space the surface may be taken sufficiently large so that $\mathbf{c}$ is zero over the entire integration. Hence Eq. (1-16) is zero, and the proof is complete.

We have thus proved that if the source density s and the circulation density $\mathbf{c}$ of a vector field $\mathbf{V}$ are given everywhere, then a solution for $\mathbf{V}$ can be derived from a *scalar potential* ϕ and a *vector potential* $\mathbf{A}$. The potentials ϕ and $\mathbf{A}$ are expressed as integrals over the source and circulation densities.

It can be proved that this system of solutions is unique if the sources are bounded in space, i.e., there are no sources at infinity, and thus the fields themselves vanish at sufficiently large distance from the sources.

Suppose that there are two functions, $\mathbf{V}_1$ and $\mathbf{V}_2$, which satisfy Eqs. (1–1) and (1–2). Their difference, the function $\mathbf{W} = \mathbf{V}_1 - \mathbf{V}_2$, obeys the conditions

$$\nabla \cdot \mathbf{W} = 0, \tag{1–17}$$

$$\nabla \times \mathbf{W} = 0, \tag{1–18}$$

at every point in space and is zero at infinity. If we now show that $\mathbf{W}$ vanishes everywhere, we shall have proved that for finite sources there is only one solution for Eqs. (1–1) and (1–2). To prove this we note that if Eq. (1–18) is satisfied we can always put

$$\mathbf{W} = -\nabla\psi \tag{1–19}$$

and, from Eq. (1–17),

$$\nabla^2\psi = 0 \tag{1–20}$$

everywhere. If we apply Gauss's divergence theorem to the vector $\psi\nabla\psi$, we obtain

$$\int \psi\nabla\psi \cdot d\mathbf{S} = \int [\psi\nabla^2\psi + (\nabla\psi)^2]\, dv. \tag{1–21}$$

The left side vanishes if the boundary is taken at sufficiently large distance from the sources, since ψ tends to zero at least as $1/r$, and the first term on the right is identically zero because of Eq. (1–20). Therefore Eq. (1–21) reduces to

$$\int (\nabla\psi)^2\, dv = \int (\mathbf{W})^2\, dv = 0, \tag{1–22}$$

and hence $\mathbf{W} = \mathbf{V}_1 - \mathbf{V}_2 = 0$ everywhere. Thus $\mathbf{V}$ as given by Eq. (1–3) is unique.

We have gone into this formal proof in great detail not only because the theorems are of fundamental importance but also because the methods are of general usefulness throughout the study of electromagnetic fields. For convenience, let us summarize the results obtained:

(a) If the source density s and the circulation density $\mathbf{c}$ of a vector field $\mathbf{V}$ are given for a finite region of space and there are no sources at infinity, then $\mathbf{V}$ is uniquely defined.

(b) If $\mathbf{V}$ has sources s but no circulation density $\mathbf{c}$, $\mathbf{V}$ is derivable from a scalar potential ϕ.

(c) If $\mathbf{V}$ has circulation density $\mathbf{c}$ but no sources s, $\mathbf{V}$ is derivable from a vector potential.

(d) $\mathbf{V}$ is always derivable from a scalar and a vector potential.

(e) At points in space where s and $\mathbf{c}$ vanish $\mathbf{V}$ is derivable from a scalar potential ϕ for which $\nabla^2\phi = 0$, or from a vector potential $\mathbf{A}$ for which $\nabla \times \nabla \times \mathbf{A} = 0$. We may add that at such points the field is said to be *harmonic*.

(f) If s and $\mathbf{c}$ are identically zero everywhere, $\mathbf{V}$ vanishes everywhere.

(g) The unique solution for $\mathbf{V}$ in terms of s and $\mathbf{c}$ is given by means of the potentials as expressed by the integrals (1–4) and (1–5).

(h) We have established a systematic notation for source and field coordinates. If we add the convention that the vector $\mathbf{r}$ points *from* source *to* field point we may extend our list of useful mathematical relations:

$$\nabla^2(1/r) = -4\pi\,\delta(r),$$

$$\nabla[g(r)] = -\nabla'[g(r)],$$

$$\nabla r = -\nabla' r = \mathbf{r}/r,$$

$$\nabla \cdot \mathbf{r} = +3.$$

These properties of general vector fields will be indispensable in the physical considerations which follow. We shall have a consistent field theory representing the empirical laws of electricity and magnetism when we have written these laws as a set of equations giving the source and circulation densities, i.e., the divergence and curl, of the field vectors representing the electromagnetic fields. This is the fundamental program of classical electromagnetic theory.

1–2 The electric field. We shall first consider the electrostatic field in vacuum. The electric field is defined in terms of the force produced on a test charge q by the equation

$$\operatorname*{limit}_{q \to 0} \frac{\mathbf{F}}{q} = \mathbf{E}, \tag{1–23}$$

where $\mathbf{F}$ is the force (newtons) on the test charge q (coulombs). The definition is entirely independent of the system of units, but in the mks system* the electric field $\mathbf{E}$ defined by this equation is in volts per meter. The limit $q \to 0$ is introduced in order that the test charge will not influence the behavior of the sources of the field, which will then be independent of the presence of the test body. The requirement that the test charge be vanishingly small compared with all sources of the field raises a fundamental difficulty, since the finite magnitude of the electronic charge does not permit the limit $q \to 0$ to be carried out experimentally. This restriction therefore limits the practical validity of the definition to cases where the sources producing the field are equivalent to a large number of electronic charges. Definition (1–23) is thus entirely suitable only for macroscopic phenomena, and we shall have to exercise great care in applying it to the treatment of the elementary charges of which matter is actually composed. For microscopic processes the field cannot be defined "operationally" in terms of its effect; it must be *described* in terms of its sources, *assuming* that the macroscopic laws of the field sources are still valid.

---

* See Appendix I for a discussion of units and the relations between units of various systems.

1-3 Coulomb's law. The experimentally established law for the force between two point charges *in vacuo* was originally formulated as an action at a distance:

$$\mathbf{F}_2 = \frac{q_1 q_2 \mathbf{r}}{4\pi\epsilon_0 r^3} = -\frac{q_1 q_2}{4\pi\epsilon_0}\nabla\left(\frac{1}{r}\right). \tag{1-24}$$

Here $\mathbf{F}_2$ is the force on charge q_2 due to the presence of charge q_1, and $\mathbf{r}$ is the radius vector position of charge q_2 measured from an origin located at charge q_1; ϵ_0 is a constant, $10^7/4\pi c^2$ farads/meter in this system of units; c is the experimental value of the velocity of propagation of plane electromagnetic waves in free space (see Appendix I), and all distances are measured in meters. In the mathematical identity on the right the gradient operator acts on the coordinates of the charge q_2. The law applies equally to positive and negative charges, and indicates that like charges repel, unlike charges attract. If the test charge in Eq. (1-23) is assumed positive, comparison with Eq. (1-24) yields immediately

$$\mathbf{E} = \frac{q}{4\pi\epsilon_0}\frac{\mathbf{r}}{r^3} = -\frac{q}{4\pi\epsilon_0}\nabla\left(\frac{1}{r}\right), \tag{1-25}$$

giving the electric field $\mathbf{E}$ at the position $\mathbf{r}$ due to a charge q at the origin of the radius vector. Here q corresponds to q_1 of Eq. (1-24).

We can prove Gauss's flux theorem

$$\int_S \mathbf{E}\cdot d\mathbf{S} = \frac{q}{\epsilon_0} \tag{1-26}$$

as a direct consequence of Coulomb's law: Consider an element of surface $d\mathbf{S}$, expressed as a vector directed along the outward normal of the element as shown in Fig. 1-1, at a distance r from a charge q at a point

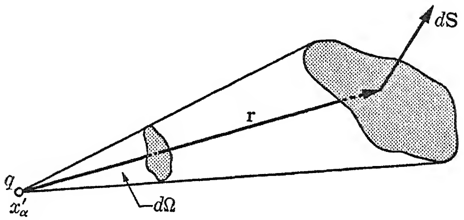


FIG. 1-1 Elements of surface and solid angle contributing to the total electric flux of Gauss's theorem.

AD AGRICULTURAL
19781

x'_α. By taking the scalar product of both sides of Eq. (1–25) we secure

$$\mathbf{E} \cdot d\mathbf{S} = \frac{1}{4\pi\epsilon_0} \frac{q}{r^3} \mathbf{r} \cdot d\mathbf{S} = \frac{q}{4\pi\epsilon_0} d\Omega.$$

Since the integral of $d\Omega$ over a closed surface which includes the point x'_α is just 4π, Gauss's theorem follows immediately. The principle of superposition enables us to sum the separate fields of any number of point charges, so that q of Eq. (1–26) is the total charge inside the boundary surface S.

If we apply the divergence theorem to $\mathbf{E}$,

$$\int_S \mathbf{E} \cdot d\mathbf{S} = \int \nabla \cdot \mathbf{E}\, dv,$$

and make use of the fact that the total volume integral of the charge density ρ is simply the total charge q, the application of the flux theorem enables us to put Eq. (1–25) into the form

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}. \tag{1–27}$$

Here ρ is the charge per unit volume at the point where the electric field is $\mathbf{E}$. Since the curl of the gradient of a scalar is zero, it further follows from Eq. (1–25) and the principle of superposition that

$$\nabla \times \mathbf{E} = 0. \tag{1–28}$$

The electrostatic field is thus irrotational. That the electrostatic field is completely defined by a charge distribution then follows from the theorem that a vector field is uniquely determined by the curl and the divergence of the field.

It is instructive to note that Eqs. (1–27) and (1–28) follow directly from Coulomb's law in the form

$$\mathbf{E}(x_\alpha) = \frac{1}{4\pi\epsilon_0} \int \rho(x'_\alpha) \frac{\mathbf{r}}{r^3} dv' \tag{1–29}$$

$$= -\frac{1}{4\pi\epsilon_0} \int \rho \nabla \left(\frac{1}{r}\right) dv'. \tag{1–30}$$

It follows that

$$\nabla \cdot \mathbf{E} = -\frac{1}{4\pi\epsilon_0} \int \rho \nabla^2 \left(\frac{1}{r}\right) dv' = \frac{1}{\epsilon_0} \int \rho(x'_\alpha)\, \delta(r)\, dv' = \frac{\rho(x_\alpha)}{\epsilon_0}. \tag{1–27'}$$

GIFT

Also

$$\nabla \times \mathbf{E} = -\frac{1}{4\pi\epsilon_0}\int \rho \nabla \times \nabla\left(\frac{1}{r}\right) dv' = 0 \qquad (1\text{–}28')$$

since the curl of a vector expressible as a gradient vanishes identically.

1–4 The electrostatic potential. Since the static field is irrotational, it may be expressed as the gradient of a scalar function. We may define an electrostatic potential ϕ by the equation

$$\mathbf{E} = -\nabla\phi. \qquad (1\text{–}31)$$

In Cartesian coordinates the components of the field parallel to the x_α axes respectively are given by

$$E_\alpha = -\frac{\partial\phi}{\partial x_\alpha}. \qquad (1\text{–}32)$$

The application of the general vector relation known as Stokes' theorem,

$$\int (\nabla \times \mathbf{E}) \cdot d\mathbf{S} = \oint \mathbf{E} \cdot d\mathbf{l} \qquad (1\text{–}33)$$

where $d\mathbf{l}$ is the infinitesimal vector length tangent to a closed path of integration, leads to

$$\oint \mathbf{E} \cdot d\mathbf{l} = 0, \qquad (1\text{–}34)$$

since the curl of $\mathbf{E}$ is everywhere zero. This shows that the electrostatic field is a conservative field: no work is done on a test charge if it is moved around a closed path in the field. Since the work done in moving a test charge from one point to another is independent of the path, we can uniquely define the work necessary to carry a unit charge from an infinite distance to a given point as the potential of that point. If one considers fields of less than three dimensions, i.e., sources extending to infinity in one or more directions, this definition will lead to difficulties and a point other than infinity must be taken as a reference point. So long as only finite sources are considered, however, this definition of potential is both adequate and convenient.

The substitution of Eq. (1–31) in Eq. (1–27) leads at once to Poisson's equation

$$\nabla^2\phi = -\rho/\epsilon_0, \qquad (1\text{–}35)$$

and in a region of zero charge density to Laplace's equation,

$$\nabla^2\phi = 0. \qquad (1\text{–}36)$$

The fundamental problem of electrostatics is to determine solutions of Poisson's equation appropriate to the conditions under particular consideration.

1-5 The potential in terms of charge distribution. The electrostatic potential at a given point was defined by Eq. (1–31) in terms of the electric field at that point. Since the source density of the electrostatic field is just ρ/ϵ_0, we know from Eq. (1–4) that the potential is

$$\phi(x_\alpha) = \frac{1}{4\pi\epsilon_0}\int \frac{\rho(x'_\alpha)\,dv'}{r(x_\alpha,x'_\alpha)}$$

in terms of the charge density at all points in space. Field theory, however, permits us to find a solution even if $\rho(x'_\alpha)$ is known only within an arbitrary surface S; the effect of the other sources is then replaced by the knowledge of the *boundary values* of the potentials or their derivatives on the surface S.

To obtain an explicit expression for $\phi(x_\alpha)$ in terms of ρ within S and ϕ and $\nabla\phi$ on S, we make use of Green's theorem, which states

$$\int (\phi\nabla^2\psi - \psi\nabla^2\phi)\,dv = \int (\phi\nabla\psi - \psi\nabla\phi)\cdot d\mathbf{S}. \tag{1–37}$$

Here ϕ and ψ are scalar functions of position that are continuous with continuous first and second derivatives in the region of integration and on its boundary. The integration will be taken over the source coordinates, x'_α, so that the field point x_α is simply a parameter. Let ϕ be the electrostatic potential defined in Eq. (1–31), and let $\psi = 1/r$, the point source solution of Laplace's equation. Since $\nabla(1/r) = -\mathbf{r}/r^3$ and we have from Eq. (1–7) that $\nabla^2(1/r) = -4\pi\,\delta(r)$, substitution of these expressions into Eq. (1–37) gives

$$-4\pi\int\left[\phi(x'_\alpha)\,\delta(r) - \frac{1}{r}\frac{\rho(x'_\alpha)}{4\pi\epsilon_0}\right]dv' = \int\left[-\phi\frac{\mathbf{r}}{r^3} - \frac{\nabla\phi}{r}\right]\cdot d\mathbf{S}.$$

Hence

$$\phi(x_\alpha) = \frac{1}{4\pi\epsilon_0}\int\frac{\rho(x'_\alpha)}{r}\,dv' + \frac{1}{4\pi}\int\left(\phi\frac{\mathbf{r}}{r^3} + \frac{\nabla\phi}{r}\right)\cdot d\mathbf{S}. \tag{1–38}$$

Note that we have made use of the functional properties of the δ-function. Clearly, Eq. (1–38) could also have been obtained without the aid of the δ-function by a limiting process. Such a process was actually implicit in the derivation of Eq. (1–7), however, so that it is unnecessary here.

The first integral of Eq. (1–38) is simply the contribution of the volume charge distribution within v. Note that the distance r is a function of the coordinates both of the point of observation, x_α, and of the point of integration, x'_α.

The surface integrals of Eq. (1–38) summarize the effect of any charge distribution outside the region v, and thus not contained in the first term. We therefore conclude that the potential at any point within S is uniquely determined by the charge distribution within S and by the values of ϕ and the normal component of $\nabla\phi$ at all points on the surface S. In particular, the potential within a charge-free volume is uniquely determined by the potential and its normal derivative over the surface enclosing the volume. What we have shown here is that knowledge of the potential and its normal derivative over the surface is sufficient to determine the potential inside, but we have not shown that both these pieces of information are necessary. We shall see later that it is, in fact, sufficient in a charge-free region to have either the potential *or* its normal derivative over a surface in order to determine the potential at every point within the surface except for an arbitrary additive constant. The reason is that ϕ and $\nabla\phi$ may not be independently specified over the surface, since ϕ must be a solution of Laplace's equation.

If the surface S is extended to include all charges in space, and arbitrarily expanded away from such charges, then the second and third terms of Eq. (1–38) vanish. This is true because the integrands involving ϕ and $\nabla\phi$ decrease at large distances at least as the inverse third power of r, while the surface of integration increases only as r^2. In later sections we shall be led to a physical interpretation of these surface integrals as equivalent to a charge and dipole distribution on the surface S. We shall therefore be able to conclude that the potential can be calculated by the direct superposition of the individual potentials of all the volume charge distribution, but that we can, if we wish, replace any part of the distribution by an equivalent surface charge layer and dipole layer distribution.

The volume term of Eq. (1–38) can be looked on as being a particular integral of Poisson's equation, while the surface terms are complementary integrals of the differential equation in the sense that they are general solutions of the homogeneous equation, i.e., Laplace's equation.

1–6 Field singularities. We have written the solution of the potential problem as a sum of boundary contributions and a volume integral extending over the source charges. These volume integrals will not lead to singular values of the potentials (or of the fields) if the charge density is bounded. If, on the other hand, the charges are considered to be surface, line, or point charges, then singularities will result as shown in Table 1–1. Note that if either surface or line charges are infinite in spatial extent (i.e., the fields are considered one- or two-dimensional) the potential cannot be referred to infinity. Although these singularities do not actually exist in nature, the fields that do occur are often indistinguishable, over much of the region concerned, from those of simple geometrical configurations. The idealizations of real charges as points, lines, and surfaces not only permit

TABLE 1–1

FIELD SINGULARITIES

Type of charge distribution	Behavior of potential near distribution	Behavior of field near distribution
Surface	r	Constant
Line	$\log r$	r^{-1}
Point	r^{-1}	r^{-2}
Point 2^n pole	r^{-n-1}	r^{-n-2}

great mathematical simplicity, they also give rise to convenient physical concepts for the description and representation of actual fields. For these reasons we shall now discuss in more detail the nature of potentials corresponding to such sources.

1–7 Clusters of point charges. The potential ϕ at the point x_α, due to the charge q located at the point x'_α, is given by

$$\phi = \frac{1}{4\pi\epsilon_0}\frac{q}{r} = \frac{1}{4\pi\epsilon_0}\left\{\frac{q}{\sqrt{(x-x')^2+(y-y')^2+(z-z')^2}}\right\}$$

$$= \frac{1}{4\pi\epsilon_0}\left\{\frac{q}{[\Sigma(x_\alpha - x'_\alpha)^2]^{1/2}}\right\}. \tag{1–39}$$

It has a first-order singularity at the point x'_α corresponding to $r = 0$. Singularities of higher order can be generated by superposing on this potential a potential corresponding to an equal charge of opposite sign, displaced a distance $\Delta x'$ from the original charge. This process is equivalent to differentiating Eq. (1–39) with respect to x'_α. We have noted in Eq. (1–15), however, that differentiation of a function of the relative coordinates only with respect to x'_α gives the same result as differentiation with respect to x_α except for sign, and at the field point Laplace's equation holds. Since the derivative of a solution of Laplace's equation is also a solution, the process of differentiation with respect to the source point as physically described above will generate new solutions with successively higher order singularities near $r = 0$. Such potentials are called multipole potentials.

For a single differentiation, we obtain

$$\phi^{(2)} = \frac{\partial \phi}{\partial x'}\Delta x' = \frac{q\,\Delta x'(x-x')}{4\pi\epsilon_0 r^3} = (q\,\Delta x')\,\frac{\cos\theta}{4\pi\epsilon_0 r^2}. \tag{1–40}$$

If we let

$$q\,\Delta\mathbf{x}' = \mathbf{p}^{(1)} \tag{1–41}$$

be the dipole moment of the distribution (positive from $-q$ to $+q$ as indicated in Fig. 1-2), we can write this potential as

$$\phi^{(2)} = \frac{1}{4\pi\epsilon_0}\mathbf{p}^{(1)}\cdot\nabla'\left(\frac{1}{r}\right) = -\frac{1}{4\pi\epsilon_0}\mathbf{p}^{(1)}\cdot\nabla\left(\frac{1}{r}\right) = \frac{1}{4\pi\epsilon_0}\frac{\mathbf{p}^{(1)}\cdot\mathbf{r}}{r^3}. \quad (1\text{-}42)$$

(The sign conventions for $\mathbf{r}$ and ∇ have been discussed in Sec. 1-1.) This solution is the dipole potential.

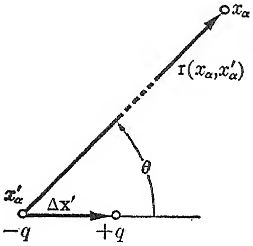


FIG. 1-2 The generation of an electric dipole.

The potential distribution, and consequently the fields of higher moments of the charge, or multipoles, can be generated by the same method of geometrical construction. For example, the potential field of a 2^{n+1} pole is generated by taking the potential of a 2^n pole and subtracting from it the potential of another 2^n pole that is displaced infinitesimally in an arbitrary direction (or by superposing the potential of the displaced 2^n pole with opposite sign).

The general form of the potential corresponding to a 2^n pole is thus

$$\phi^{(2^n)} = \frac{p^{(n)}}{4\pi\epsilon_0 n!}\frac{\partial^n}{\partial x'_\alpha \partial x'_\beta \cdots}\left(\frac{1}{r}\right) = \frac{(-1)^n p^{(n)}}{4\pi\epsilon_0 n!}\frac{\partial^n}{\partial x_\alpha \partial x_\beta \cdots}\left(\frac{1}{r}\right) \quad (1\text{-}43)$$

in terms of the multipole moment $p^{(n)}$, defined by the recurrence relation $p^{(n)} = np^{(n-1)}\,\Delta x'_n$, where $\Delta x'_n$ is the displacement leading to the 2^n pole. The displacement need not be along coordinate axes, but the derivative corresponding to an oblique displacement can be written as a sum of derivatives with respect to x, y, z, having the direction cosines of the displacement as coefficients. A few examples of simple multipoles are shown in Fig. 1-3. In the special case where all the displacements are in one direction the problem has axial symmetry and only one angle is needed to specify the position of the field point. For a linear 2^n pole,

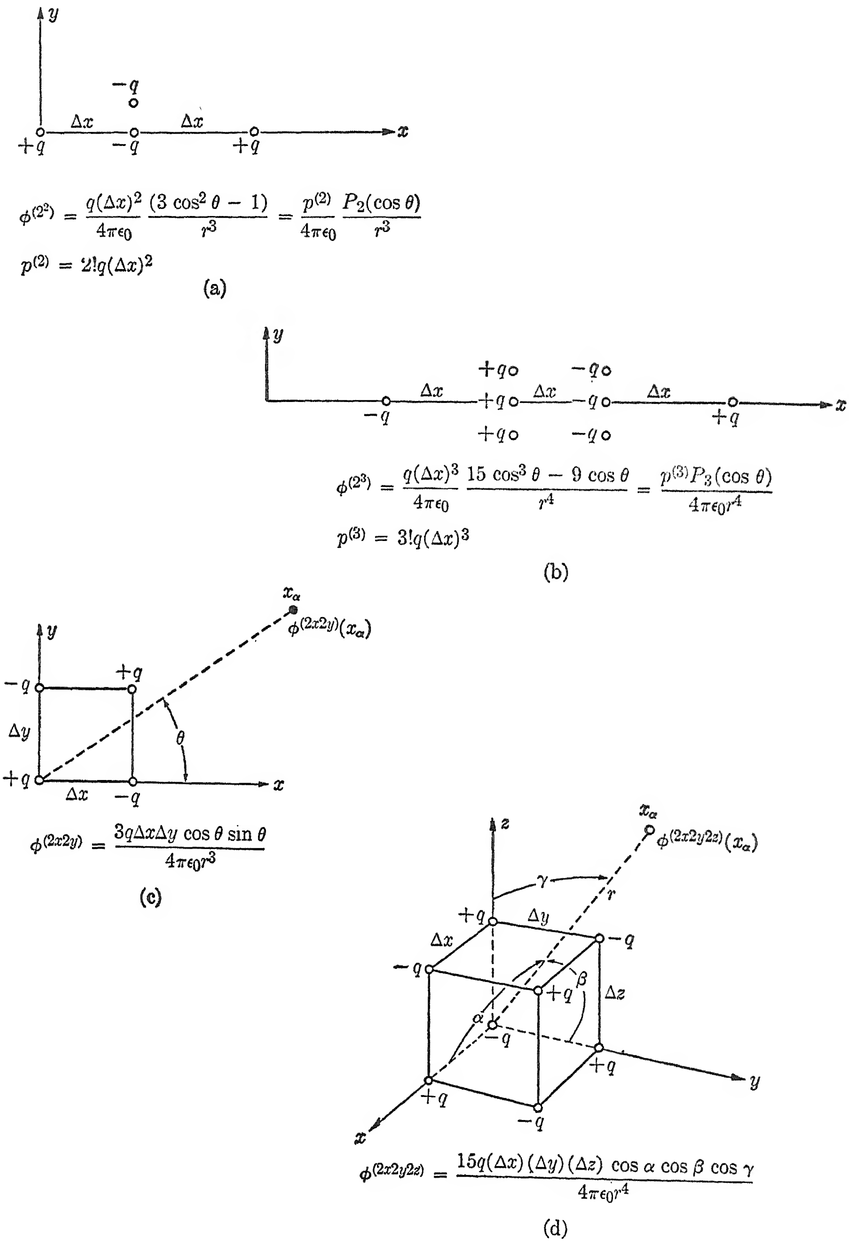
$$\phi^{(2^n)}(x,y,z) = \frac{p^{(n)}}{4\pi\epsilon_0 n!}\frac{\partial^n}{\partial x'^n}\left(\frac{1}{r}\right) = \frac{p^{(n)}}{4\pi\epsilon_0}\frac{P_n(\cos\theta)}{r^{n+1}}, \quad (1\text{-}44)$$

where $P_n(\cos\theta)$ is the Legendre polynomial of order n which may be defined by the relation

$$\frac{P_n(\cos\theta)}{r^{n+1}} = \frac{(-1)^n}{n!}\frac{\partial^n}{\partial x^n}\left(\frac{1}{r}\right), \quad (1\text{-}45)$$

in which θ is the angle between x and r.

Equation (1-43) can be recognized as the nth term of a general Taylor expansion of $1/r$ in terms of the source coordinates. The coefficients are the multipole moments as defined in terms of the specific charge distribu-


y
x
+q
−q
+q
−q
Δx
Δy
Δz
θ
α
β
γ
r
z
x_α
$\phi^{(2x2y)}(x_\alpha)$
$\phi^{(2x2y2z)}(x_\alpha)$

$$\phi^{(2^2)} = \frac{q(\Delta x)^2}{4\pi\epsilon_0}\frac{(3\cos^2\theta - 1)}{r^3} = \frac{p^{(2)}}{4\pi\epsilon_0}\frac{P_2(\cos\theta)}{r^3}$$

$$p^{(2)} = 2!q(\Delta x)^2$$

(a)

$$\phi^{(2^3)} = \frac{q(\Delta x)^3}{4\pi\epsilon_0}\frac{15\cos^3\theta - 9\cos\theta}{r^4} = \frac{p^{(3)}P_3(\cos\theta)}{4\pi\epsilon_0 r^4}$$

$$p^{(3)} = 3!q(\Delta x)^3$$

(b)

$$\phi^{(2x2y)} = \frac{3q\Delta x\Delta y\cos\theta\sin\theta}{4\pi\epsilon_0 r^3}$$

(c)

$$\phi^{(2x2y2z)} = \frac{15q(\Delta x)(\Delta y)(\Delta z)\cos\alpha\cos\beta\cos\gamma}{4\pi\epsilon_0 r^4}$$

(d)

FIG. 1-3 Examples of multipoles: (a) linear quadrupole; (b) linear octupole; (c) two-dimensional quadrupole; (d) three-dimensional octupole.

tion referred to above. We shall now show that the potential of an arbitrary charge distribution $\rho(x'_\alpha)$ of finite extent can at large distances always be expressed as the sum of multipole potentials where the coefficients are certain integrals (moments) of the charge distribution.

To facilitate the proof of this statement, we shall choose the origin of coordinates in or near the charge distribution. Let R be the distance from the origin O to the field point, i.e., let the components of $\mathbf{R}$ be x_α. We may then expand $1/r$ in powers of x'_α about O, assuming that $a/R \ll 1$, where a is a limiting dimension of the bounded distribution of charge. By Taylor's theorem,

$$\frac{1}{r} = \frac{1}{R} + x'_\alpha \left[\frac{\partial}{\partial x'_\alpha}\left(\frac{1}{r}\right)\right]_{r=R} + \frac{1}{2!} x'_\alpha x'_\beta \left[\frac{\partial^2}{\partial x'_\alpha \partial x'_\beta}\left(\frac{1}{r}\right)\right]_{r=R} + \cdots. \qquad (1\text{–}46)$$

In Eq. (1–46) we are employing the "summation convention" which we shall continue to use throughout: when indices are repeated in the same term summation over these indices is implied. Upon substituting Eq. (1–46) in the general potential, we obtain the "multipole expansion":

$$\begin{aligned}\phi(x_\alpha) &= \frac{1}{4\pi\epsilon_0}\int \frac{\rho(x'_\alpha)}{r}\,dv' \\ &= \frac{1}{4\pi\epsilon_0}\left\{\frac{1}{R}\int \rho\,dv' + \left[\frac{\partial}{\partial x'_\alpha}\left(\frac{1}{r}\right)\right]_{r=R}\int x'_\alpha \rho\,dv' \right. \\ &\quad \left. + \frac{1}{2!}\left[\frac{\partial^2}{\partial x'_\alpha \partial x'_\beta}\left(\frac{1}{r}\right)\right]_{r=R}\int x'_\alpha x'_\beta \rho\,dv' + \cdots\right\}. \qquad (1\text{–}47)\end{aligned}$$

The coefficients represent the moments of the charge distribution: $\int \rho\,dv'$ is the total charge q, $\int x'_\alpha \rho\,dv'$ is the α-component of the dipole moment, etc. The radial and angular dependence of each term in Eq. (1–47) is clearly identical with that given by Eq. (1–43).

Note that the words "dipole," "quadrupole," etc., are being used in two ways: first to describe a specific charge distribution, and secondly to designate moments of an arbitrary charge distribution. Both physical quantities give rise to the same potential distribution.

The ratio of the magnitudes of successive terms in the multipole expansion, Eq. (1–47), is of order a/R, where a is a parameter characteristic of the size of the charge distribution. Hence if $R \gg a$ the potential of an arbitrary distribution may be very conveniently replaced by the potentials of the moments of the distribution.

1–8 Dipole interactions. The energy of position of an already created dipole in an electrostatic field is given by

$$U = -\mathbf{p} \cdot \mathbf{E}. \qquad (1\text{–}48)$$

This expression follows directly from the torque relation of a dipole,

$$\mathbf{L} = \mathbf{p} \times \mathbf{E}, \tag{1-49}$$

in a way indicated in the problems. Here $\mathbf{L}$ is the torque exerted in such a direction as to rotate the positive end of the dipole toward the direction of the field. The force acting on a dipole in an inhomogeneous electric field, when the relative orientation of the dipole and field is not free to change, is given by*

$$\mathbf{F} = -\nabla U = -\nabla(-\mathbf{p} \cdot \mathbf{E}) = (\mathbf{p} \cdot \nabla)\mathbf{E}. \tag{1-50}$$

Note that this force vanishes for uniform fields, as would be expected from symmetry arguments.

Let us now consider the interaction force and energy between two dipoles, such as those shown in Fig. 1-4, whose moment vectors are oriented in space at an arbitrary angle with each other. On combining the force equation above with the potential equation (1-42) and the field equation (1-31), we have for the force $\mathbf{F}_1$ on dipole 1 in the field of dipole 2 and, conversely, for the force $\mathbf{F}_2$ on dipole 2 due to dipole 1,

$$\phi_1 = \frac{-1}{4\pi\epsilon_0}\,\mathbf{p}_2 \cdot \nabla_1\left(\frac{1}{r}\right); \quad \mathbf{E}_1 = -\nabla_1\phi_1 = \frac{1}{4\pi\epsilon_0}\,\nabla_1\left[\mathbf{p}_2 \cdot \nabla_1\left(\frac{1}{r}\right)\right];$$

$$\mathbf{F}_1 = (\mathbf{p}_1 \cdot \nabla_1)\mathbf{E}_1 = \frac{1}{4\pi\epsilon_0}\,(\mathbf{p}_1 \cdot \nabla_1)\left\{\nabla_1\left[\mathbf{p}_2 \cdot \nabla_1\left(\frac{1}{r}\right)\right]\right\}; \tag{1-51}$$

$$\mathbf{F}_2 = \frac{1}{4\pi\epsilon_0}\,(\mathbf{p}_2 \cdot \nabla_2)\left\{\nabla_2\left[\mathbf{p}_1 \cdot \nabla_2\left(\frac{1}{r}\right)\right]\right\} = -\mathbf{F}_1.$$

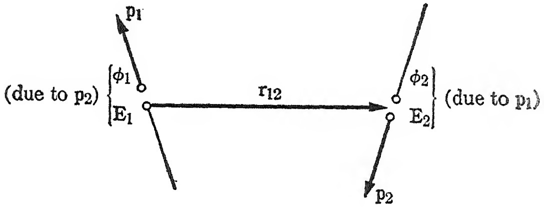


FIG. 1-4 Interaction of two dipoles.

The interaction energy between the two dipoles may be obtained by inserting the field $\mathbf{E}_1$ above into the expression for the energy U. We

---

[*] Since $\nabla(\mathbf{p} \cdot \mathbf{E}) = \mathbf{p} \times (\nabla \times \mathbf{E}) + (\mathbf{p} \cdot \nabla)\mathbf{E} = (\mathbf{p} \cdot \nabla)\mathbf{E}$ *if* $\nabla \times \mathbf{E} = 0$. If curl $\mathbf{E}$ is not equal to zero, an additional term is obtained.

find for the energy U_{12} of dipole 1 in the field of dipole 2, and conversely for U_{21}:

$$U_{12} = -\mathbf{p}_1 \cdot \mathbf{E}_1 = -\frac{(\mathbf{p}_1 \cdot \nabla_1)}{4\pi\epsilon_0}\left[\mathbf{p}_2 \cdot \nabla_1 \left(\frac{1}{r}\right)\right],$$

$$U_{12} = \frac{1}{4\pi\epsilon_0}\left[\frac{\mathbf{p}_1 \cdot \mathbf{p}_2}{r^3} - \frac{3}{r^5}(\mathbf{p}_1 \cdot \mathbf{r})(\mathbf{p}_2 \cdot \mathbf{r})\right], \tag{1–52}$$

$$U_{21} = U_{12}.$$

This is the general expression for the interaction energy of two dipoles.

1–9 Surface singularities. Surface singularities of the second order or dipole form are of particular interest in both electrostatics and magnetostatics. Let us consider a double layer charge arrangement with a dipole moment per unit area designated by $\boldsymbol{\tau}$. The potential arising from such a distribution is given by

$$\phi = \frac{1}{4\pi\epsilon_0}\int \frac{\boldsymbol{\tau} \cdot \mathbf{r}}{r^3}\, dS. \tag{1–53}$$

This expression reduces, in the case when $\boldsymbol{\tau}$ is uniform and normal to the surface over the dipole sheet, to

$$\phi = \frac{|\boldsymbol{\tau}|}{4\pi\epsilon_0}\int \frac{\mathbf{r} \cdot d\mathbf{S}}{r^3} = \frac{|\boldsymbol{\tau}|}{4\pi\epsilon_0}\,\Omega. \tag{1–54}$$

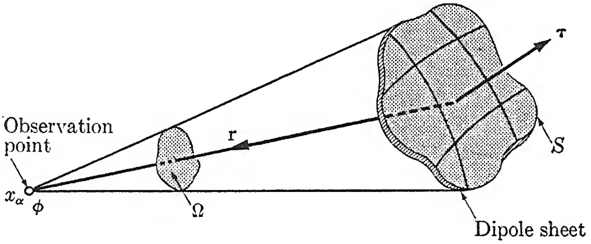
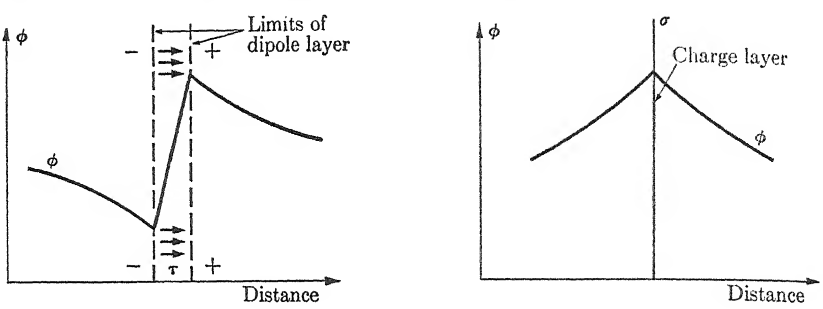


FIG. 1–5 Potential due to dipole layer.

Here Ω is the solid angle subtended by the dipole sheet at the point of observation, as in Fig. 1–5. The solid angle subtended by a nonclosed surface jumps discontinuously by 4π as the point of observation crosses the surface. This means that in the ideal case of an infinitely thin dipole charge layer the potential function will have a discontinuity of magnitude $|\boldsymbol{\tau}|/\epsilon_0$, but it will have a continuous derivative at the dipole sheet.



(a) (b)

FIG. 1–6 Behavior of the potential at a dipole layer, (a), and at a layer of charge, (b).

On the other hand, a simple surface charge layer will not result in a discontinuity in potential, but will produce a discontinuity in the normal derivative of the potential, the magnitude of discontinuity being σ/ϵ_0, where σ is the surface charge density of the layer. A comparison between the two cases is shown in Fig. 1–6(a) and (b). Since surface charge layers and dipole layers enable us to introduce arbitrary discontinuities in the potential and its derivatives at a particular surface, we can make the potential vanish outside a given volume by surrounding the volume with a suitably chosen charge layer and dipole layer. This is a further explanation of the significance of the surface terms in Eq. (1–38), which was derived from Green's theorem. These terms, when ϕ and $\nabla\phi$ are properly evaluated on the surface in terms of τ and σ, are precisely those necessary to cancel the field of those charges inside the surface S in the region outside of S. This can be seen by writing Eq. (1–38) as

$$\phi = \frac{1}{4\pi\epsilon_0}\left(\int \frac{\rho\,dv}{r} + \int \tau\frac{\mathbf{r}\cdot d\mathbf{S}}{r^3} + \int \sigma\frac{dS}{r}\right), \tag{1–55}$$

where $\tau = \epsilon_0\phi$ and σ is ϵ_0 times the normal derivative of ϕ.

As an example of a combined surface charge and dipole layer that will just cancel the field outside a given surface, yet leave the field inside the surface unchanged, consider a point charge q located at the point $R = 0$, and the surface $R = a$ surrounding this charge. If we place a surface charge density $\sigma = -q/4\pi a^2$ per unit area on the sphere $R = a$, it will give rise to a potential:

$$\phi_\sigma = -\frac{q}{4\pi\epsilon_0 a}, \quad \text{for } R < a,$$

$$\phi_\sigma = -\frac{q}{4\pi\epsilon_0 R}, \quad \text{for } R > a.$$

If, in addition, a surface dipole layer of moment $\tau = q\mathbf{R}/4\pi aR$ per unit area is placed on the sphere, it will make a contribution

$$\phi_\tau = \frac{q}{4\pi\epsilon_0 a}, \quad R < a,$$

$$\phi_\tau = 0, \quad R > a.$$

The potential of the original charge q is $\phi_0 = q/4\pi\epsilon_0 R$ for all R. The total of all these potentials is

$$\phi = \phi_0 + \phi_\sigma + \phi_\tau = \frac{q}{4\pi\epsilon_0 R}, \quad \text{for } R < a,$$

$$\phi = 0, \quad \text{for } R > a.$$

The electric field produced by a dipole layer of area S as shown in Fig. 1–7 can be derived as follows. Consider a change in the potential corresponding to a displacement of the point of observation x_α by a distance $d\mathbf{x}$,

$$d\phi = -\mathbf{E}\cdot d\mathbf{x}. \qquad (1\text{–}56)$$

The change in solid angle, $d\Omega$, subtended by the dipole layer at the field point is the same whether the field point moves a distance $d\mathbf{x}$ or the layer moves through $-d\mathbf{x}$. The latter case is shown in the figure. Since in this displacement an element $d\mathbf{l}$ of the boundary sweeps over an area $d\mathbf{x} \times d\mathbf{l}$, the total change in solid angle is

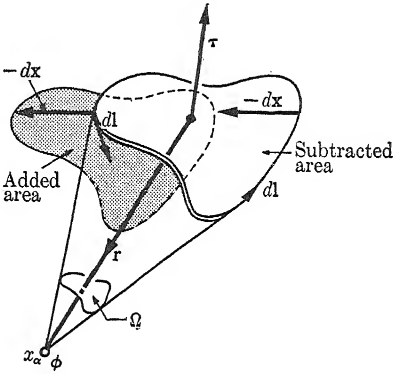


FIG. 1–7 Illustrating the derivation of the electric field produced by a dipole layer.

$$d\Omega = \oint \frac{(d\mathbf{x} \times d\mathbf{l})\cdot \mathbf{r}}{r^3} = \oint \frac{(d\mathbf{l} \times \mathbf{r})\cdot d\mathbf{x}}{r^3}. \qquad (1\text{–}57)$$

The change in potential corresponding to this change in solid angle is, from Eq. (1–54),

$$d\phi = -\frac{\tau}{4\pi\epsilon_0}\, d\Omega. \qquad (1\text{–}58)$$

(The negative sign follows from the fact that the negative side of the dipole layer is toward the observer at x_α.) Equating the two expressions for $d\phi$, we obtain

$$-\mathbf{E}\cdot d\mathbf{x} = \frac{-\tau}{4\pi\epsilon_0}\oint \frac{(d\mathbf{l} \times \mathbf{r})\cdot d\mathbf{x}}{r^3}. \qquad (1\text{–}59)$$

Since $d\mathbf{x}$ is an arbitrary displacement and this last expression holds for all possible $d\mathbf{x}$, it is permissible to write

$$\mathbf{E} = \frac{\tau}{4\pi\epsilon_0}\oint \frac{d\mathbf{l}\times\mathbf{r}}{r^3} = -\frac{\tau}{4\pi\epsilon_0}\oint d\mathbf{l}\times\nabla\left(\frac{1}{r}\right). \tag{1–60}$$

That the potential due to a dipole layer is double valued at the surface is strictly true only in the limit in which the dipole layer has zero thickness (see Fig. 1–6); for this reason the discontinuity does not actually have physical reality. Nevertheless, the method of generating nonconservative potentials by means of such discontinuities is a useful one, particularly in the theory of magnetic fields due to currents, where the corresponding potential does have a multivalued behavior completely analogous to the properties of the surface dipole moment.

1–10 Volume distributions of dipole moment. The potential due to the volume distribution of dipole moment is found by considering the dipole moment in Eq. (1–42) as a volume density and integrating over the volume. If $\mathbf{P}$ is the dipole moment per unit volume,

$$\phi = \frac{1}{4\pi\epsilon_0}\int \mathbf{P}\cdot\nabla'\left(\frac{1}{r}\right)dv'. \tag{1–61}$$

This can be changed into a form that is physically more revealing by means of Gauss's divergence theorem and the relation

$$\nabla'\cdot\left(\frac{\mathbf{P}}{r}\right) = \frac{1}{r}\nabla'\cdot\mathbf{P} + \mathbf{P}\cdot\nabla'\left(\frac{1}{r}\right). \tag{1–62}$$

We obtain

$$\begin{aligned}\phi &= \frac{1}{4\pi\epsilon_0}\left[\int \nabla'\cdot\left(\frac{\mathbf{P}}{r}\right)dv' - \int \frac{1}{r}\nabla'\cdot\mathbf{P}\,dv'\right]\\ &= \frac{1}{4\pi\epsilon_0}\left[\int \frac{\mathbf{P}\cdot d\mathbf{S}}{r} - \int \frac{\nabla'\cdot\mathbf{P}}{r}\,dv'\right].\end{aligned} \tag{1–63}$$

This expression can be interpreted as follows. The first term, a surface integral, is a potential equivalent to that of a surface charge density, while the second term is a potential equivalent to that of a volume charge density. The charge densities which have potentials equivalent to those produced by the volume polarization of a region of space are

$$\sigma_P = P_n, \quad \rho_P = -\nabla'\cdot\mathbf{P}. \tag{1–64}^*$$

---

[*] Since $\rho_P = -\nabla\cdot\mathbf{P}$ is a field equation, the prime on ∇ can be dropped without ambiguity. The prime on the ∇ is only necessary in integral expressions which relate a field quantity to an integral over a source quantity.

The relations between these surface and volume charges and the polarization can be derived from purely geometrical considerations. If, for example, we have an inhomogeneous dipole moment per unit volume, ρ_P will represent the charge density that accumulates from incomplete cancellation of the ends of the individual dipoles distributed in the volume. The quantity σ_P, on the other hand, represents the charge density on the surface produced by the lack of neighbors for the dipoles which lie with their ends on the surface. It is evident that ρ_P will vanish in a homogeneous medium; in fact, a sufficient condition for its vanishing is that the dipole moment per unit volume have a zero divergence. In general, however, the potential due to the two forms of polarization charges is

$$\phi = \frac{1}{4\pi\epsilon_0}\left[\int \frac{\sigma_P\,dS}{r} + \int \frac{\rho_P\,dv'}{r}\right]. \tag{1-65}$$

Note that the "equivalence relations" of Eq. (1-64), although derived above by means of the electric field which the respective terms produce, are actually simple geometrical quantities. We can see this formally by considering the total dipole moment $\mathbf{p}$ of a distribution. According to the "equivalence relations," we should have

$$\mathbf{p} = \int \mathbf{P}\,dv' = \int \boldsymbol{\xi}\rho_P\,dv' + \int \boldsymbol{\xi}\sigma_P\,dS = -\int \boldsymbol{\xi}(\nabla'\cdot\mathbf{P})\,dv' + \int \boldsymbol{\xi}(\mathbf{P}\cdot d\mathbf{S}), \tag{1-66}$$

where $\boldsymbol{\xi}$ is the vector whose components are x'_α. The αth component of $\mathbf{p}$ can be integrated by parts from the identity

$$\nabla'\cdot(x'_\alpha\mathbf{P}) = x'_\alpha\nabla'\cdot\mathbf{P} + P_\alpha.$$

Therefore,

$$\begin{aligned} p_\alpha = \int P_\alpha\,dv' &= -\int x'_\alpha\nabla'\cdot\mathbf{P}\,dv' + \int \nabla'\cdot(x'_\alpha\mathbf{P})\,dv' \\ &= -\int x'_\alpha\nabla'\cdot\mathbf{P}\,dv' + \int x'_\alpha\mathbf{P}\cdot d\mathbf{S}. \end{aligned}$$

Equation (1-66) is thus proved as a geometrical relationship without reference to any interaction.

SUGGESTED REFERENCES*

M. Abraham and R. Becker, *The Classical Theory of Electricity and Magnetism* (Volume 1). The standard reference on a high intermediate level, beginning with an excellent summary of vector analysis and the properties of vector fields. Gaussian units are used throughout Abraham-Becker and in most other classical texts on electrodynamics. For a comparison of units, see Appendix I of the present book.

---

* Full bibliographical description of the references listed at the end of each chapter will be found in the bibliography at the end of the book.

G. P. HARNWELL, *Principles of Electricity and Electromagnetism.* Intermediate text with technical applications.

J. H. JEANS, *Electricity and Magnetism.* A comprehensive treatment of electrostatics, although one which avoids vector notation.

W. R. SMYTHE, *Static and Dynamic Electricity.* The most comprehensive modern treatment (in English) of electrostatics and the solution of potential problems.

J. A. STRATTON, *Electromagnetic Theory.* This work is developed from Maxwell's equations as postulates, so that an account of the electrostatic field is considerably postponed. The theory of multipoles is explicit for nonorthogonal displacements.

The electrostatic field is also treated in well-known textbooks of theoretical physics:

P. G. BERGMANN, *Basic Theories of Physics: Mechanics and Electrodynamics.*

G. JOOS, *Theoretical Physics.*

L. PAGE, *Introduction to Theoretical Physics.*

J. SLATER AND N. FRANK, *Introduction to Theoretical Physics* or *Electromagnetism.*

The properties of Legendre functions are summarized by Jeans and by Smythe, for example, and are also to be found in useful form in such mathematical references as:

H. MARGENAU AND G. M. MURPHY, *The Mathematics of Physics and Chemistry.*

L. A. PIPES, *Applied Mathematics for Engineers and Physicists.*

EXERCISES

1. Show that Green's theorem, Eq. (1–37), follows from Gauss's divergence theorem.

2. We shall need the "one-dimensional" δ-function defined by

$$\int \delta(x - a)\,dx = 1,$$

$$\int f(x)\,\delta(x - a)\,dx = f(a)$$

when $x = a$ is in the interval of integration, and both integrals vanish otherwise. Show by means of the Fourier integral theorem that

$$\int_{-\infty}^{\infty} \cos kx\,dk = 2\pi\,\delta(x)$$

in the sense that both behave the same way as factors in an integrand.

3. Prove by considering the axial point on a disk that the potential undergoes no sudden change from one side to the other of a charge layer, and that the same statement holds for the normal derivative of the potential in the case of a dipole layer.

4. Functions of the type $\phi = x$, or indeed $\phi = x^2 + 2y^2 - 3z^2$, satisfy Laplace's equation at all points of space. Does this mean that such potentials have no sources? Discuss in detail the significance of such solutions, and their bearing on the uniqueness proof for potentials.

5. If the electric field of a point charge q were proportional to $qr^{-2-\delta}\hat{\mathbf{r}}$, where $\hat{\mathbf{r}}$ is a radial unit vector and $\delta \ll 1$, (a) calculate $\nabla \cdot \mathbf{E}$ and $\nabla \times \mathbf{E}$, for $r \neq 0$; (b) and if two concentric spherical shells of radii a and b were connected by a wire, with

q_a on the outer shell, prove that

$$q_b = \frac{-q_a\delta}{2(a-b)}\left\{[2b \log 2a - (a+b)\log(a+b) + (a-b)\log(a-b)]\right\} + 0(\delta^2)$$

will reside on the inner shell. (Adapted from Jeans.)

6. Show that the static potential $\phi(x,y,z)$, correct to third order in a, is equal to the mean of the potentials at the points

$$x \pm a,\ y,\ z; \quad x,\ y \pm a,\ z; \quad x,\ y,\ z \pm a.$$

7. For a finite spherically symmetric charge distribution the potential as calculated by $\phi = \int_r^\infty \mathbf{E} \cdot dr$ is

$$\phi(r) = \frac{1}{\epsilon_0}\int_r^\infty \frac{dr''}{r''^2}\int_0^{r'} \rho r'^2\, dr'.$$

By dividing the distribution into thin shells, each of which contributes constant potential at all points inside it, obtain an expression for ϕ that involves only single integrals. Prove the equality of the two expressions.

8. (a) From the relation between torque and potential energy show that $U = -\mathbf{p} \cdot \mathbf{E}$ if $\mathbf{L} = \mathbf{p} \times \mathbf{E}$ for a dipole of moment $\mathbf{p}$.

(b) Consider two coplanar electric dipoles with their centers a fixed distance apart. Show that if the angles the dipoles make with the line joining their centers are θ and θ' respectively, and if θ is held fixed,

$$\tan\theta = -\tfrac{1}{2}\tan\theta'$$

for equilibrium.

9. The differential equations of the "lines of force" are

$$\frac{dx}{E_x} = \frac{dy}{E_y} = \frac{dz}{E_z}.$$

For a dipole of moment $\mathbf{p}$ directed along the x-axis and located at the origin, find the equation $f(x,y) =$ constant that gives the lines of force in the plane $z = 0$.

10. Calculate the quadrupole moment of two concentric ring charges q and $-q$, having radii a and b.

11. Two uniform line charges, each of length $2a$, cross each other at the origin in such a way that their ends are at the points $(\pm a,0,0)$ and $(0,\pm a,0)$. Determine ϕ for points $r > a$, up to but not including terms in r^{-4}.

12. Show that the potential of a symmetrical 2^n pole generated by differentiating the point potential n times along successive directions making an angle $2\pi/n$ with one another is given by

$$\phi(r,\theta,\varphi) = \frac{\text{const.}}{r^{n+1}} P_n^n(\cos\theta)\cos n(\varphi - \varphi_0).$$

13. Show that $\mathbf{F}_1 = -\mathbf{F}_2$ in Eq. (1–51).

CHAPTER 2

BOUNDARY CONDITIONS AND RELATION OF MICROSCOPIC TO MACROSCOPIC FIELDS

The dipole moments per unit volume considered in the foregoing chapter are special examples of sources which give rise to electrostatic fields and can therefore be treated as special types of charge densities in Poisson's equation. Since such volume distributions are produced in material media by electric fields, the behavior of a medium in a field can be described in terms of its polarization, i.e., its dipole moment per unit volume. It is customary, in order to clarify the understanding of polarization, to separate the total charge that produces an electrostatic field into two parts: a true, free, movable, net charge ρ, and a bound, zero-net, polarization charge ρ_P. This division is to a certain extent arbitrary, in the sense that the polarization charge ρ_P simply represents separated charges which on the scale of observation being considered in a particular experiment are essentially inaccessible, but which would be treated as free charges on a smaller scale. If, for example, we place a piece of metal between the plates of a condenser, we can describe the resultant field between the plates either in terms of the true charges produced on the metal or in terms of an equivalent polarization of the piece of metal, depending on whether we consider the charges individually measurable. If, instead of the metal, we introduce a piece of dielectric between the condenser plates, we are forced to describe the phenomena by a polarization charge, rather than by a true charge, since it is assumed in the theory that observation shall not be made on an atomic scale. An atomic scale observation would be necessary in order to "resolve" the volume polarization into individual charges.

2–1 The displacement vector. It is seen that the distinction between ρ and ρ_P is an arbitrary one, but this arbitrariness will in no way disturb the formalism used to describe the fields produced by polarization charges. Since we have divided the sources of electric fields into these two types, the Poisson source equation becomes

$$\nabla^2\phi = -\nabla \cdot \mathbf{E} = \frac{-1}{\epsilon_0}(\rho + \rho_P). \tag{2–1}$$

The symbol ρ now denotes only the true free charge at the point where the divergence is taken. If ρ_P is expressed in terms of the divergence of the polarization $\mathbf{P}$, as given by Eq. (1-64), we obtain from Eq. (2-1)

$$\nabla \cdot \left(\mathbf{E} + \frac{\mathbf{P}}{\epsilon_0}\right) = \frac{\rho}{\epsilon_0}. \tag{2-2}$$

It is thus convenient to define an electric displacement vector $\mathbf{D}$ (measured in coulombs per square meter in the mks system) by

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}. \tag{2-3}$$

The source equations then become

$$\nabla \cdot \mathbf{D} = \rho \tag{2-4}$$

and

$$\nabla \cdot \mathbf{E} = \frac{\rho_t}{\epsilon_0}, \tag{2-5}$$

with the total charge density ρ_t written for the sum of the true and polarization charge densities. The corresponding integral relations, secured by means of Gauss's divergence theorem on integration over the volume containing all the charges, are

$$\int \mathbf{D} \cdot d\mathbf{S} = q, \tag{2-6}$$

$$\int \mathbf{E} \cdot d\mathbf{S} = \frac{q_t}{\epsilon_0}, \tag{2-7}$$

where q_t is the total charge, the sum of q and the integral of $-(\nabla \cdot \mathbf{P})$ over the volume.

It is clear that $\mathbf{D}$ represents a partial field, namely, that electric field whose sources are only the true charges. Note that the relation (2-3) between $\mathbf{D}$ and $\mathbf{E}$ is basically an additive one, the difference between $\mathbf{D}$ and $\epsilon_0\mathbf{E}$ being the polarization $\mathbf{P}$. Note also that the polarization, although defined in a purely geometrical fashion as the dipole moment per unit volume, has the properties of an electric field. The polarization field $\mathbf{P}$ is that field whose flux arises only from the polarization charges ρ_p. The solution of an actual field problem involving polarized bodies will depend on the manner in which the polarization depends on the external field. In most cases the polarization is proportional to the field, and can be expressed by an equation of the type

$$\mathbf{P} = \epsilon_0 \chi \mathbf{E}, \tag{2-8}$$

where χ is called the electric susceptibility. Such a description excludes the consideration of electrets (materials possessing a permanent dipole moment), but electrets are not ordinarily of much importance. In case we do have a simple medium whose polarization depends linearly on the imposed electric field as expressed by Eq. (2-8), then all three vectors $\mathbf{D}$,

$\mathbf{E}$, and $\mathbf{P}$ will be related by constants of proportionality:

$$\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P} = \epsilon_0(1 + \chi)\mathbf{E}. \tag{2-9}$$

We may define the specific inductive capacity (often called the dielectric constant) by

$$\kappa = 1 + \chi, \tag{2-10}$$

so that

$$\mathbf{D} = \kappa\epsilon_0\mathbf{E} \tag{2-11}$$

and

$$\mathbf{P} = \epsilon_0(\kappa - 1)\mathbf{E}. \tag{2-12}$$

As it stands, Eq. (2–8) presupposes that the medium polarizes isotropically, or that the polarization properties of the medium do not depend on the direction of the polarization. This is not the general case and, in fact, the scalar proportionality is valid only for liquids, gases, amorphous solids, and cubic crystals. In crystals of symmetry lower than cubic the relation between each of the components of the polarization vector and of the electric field vector is still linear but the constants of proportionality in the various directions may be different. This means that the relation between the components of the polarization vector and the components of the electric field vector are given by a tensor,

$$P_\alpha = \epsilon_0\chi_{\alpha\beta}E_\beta, \tag{2-13}$$

and $\mathbf{P}$ is no longer in the same direction as $\mathbf{E}$. From geometrical consideration of the fact that the magnitude of $\mathbf{P}$ is proportional to that of $\mathbf{E}$, while the angle between $\mathbf{P}$ and $\mathbf{E}$ is constant for a given orientation of the material in the field, the tensor $\chi_{\alpha\beta}$ may be shown to be symmetrical. (The details of this proof are left as a problem.) It is therefore possible to express $\chi_{\alpha\beta}$ in terms of principal coordinates by a set of only three constants, and there are at least three directions in which $\mathbf{P}$ and $\mathbf{E}$ are parallel. The symmetry of $\chi_{\alpha\beta}$ can also be deduced from energy considerations.

The case where $\mathbf{E}$ and $\mathbf{P}$ are not proportional and there is no linear relation between them will not be treated here, but the analogous case will be discussed in connection with magnetic media where nonlinearity is of more practical importance. It should be pointed out, however, that the relation we have here assumed between $\mathbf{P}$ and $\mathbf{E}$ is only a special simplification, and not a fundamental equation of the theory.

2–2 Boundary conditions. Maxwell's field equations, to be discussed later, are a set of equations whose sources are divided into accessible and inaccessible charges. To obtain a solution of Maxwell's equations the inaccessible charges must be related to the accessible charges, or to the fields produced by the accessible charges, by additional equations. The relations that evaluate the inaccessible charge sources in terms of the external fields which produce them are called the constitutive equations.

Equation (2–8) is an example of such an equation. The constitutive equations must, of course, depend on the properties of the material in which the inaccessible charges arise. While Eqs. (2–8) to (2–13) are not entirely general, they depend only on linearity and do not imply homogeneity. The susceptibility and specific inductive capacity may therefore be arbitrary functions of the coordinates. A case of much interest is one where the specific inductive capacity varies discontinuously, as at the boundary between two dielectrics.

To determine the behavior of the fields at a boundary, let us first imagine a small volume, as in Fig. 2-1, whose dimension normal to the interface, h, is smaller than its other dimensions by an order of magnitude, to be placed so that one of its large surfaces ΔA lies in medium 1, the other in medium 2, and both are parallel to the interface. This small volume can be used to derive the behavior of the normal components of the fields.

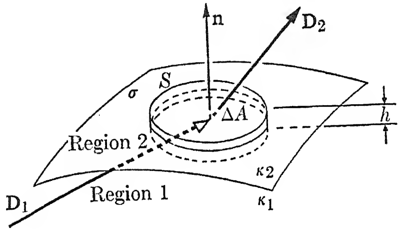


Fig. 2–1 Volume considered for determining the boundary conditions on the normal component of **D**.

Let us take the surface integral of **D** over this small volume.

$$\int \mathbf{D} \cdot d\mathbf{S} = q. \tag{2–6}$$

In the limit as the dimension h approaches zero, q approaches $\sigma\Delta A$, where σ is the true surface charge density on the interface. The contribution of the sides of the volume normal to the surface vanishes, so that Eq. (2–6) becomes

$$\mathbf{n} \cdot (\mathbf{D}_2 - \mathbf{D}_1) = \sigma, \tag{2–14}$$

where $\mathbf{n} = d\mathbf{S}/dS$ is the unit vector normal to ΔA, the top of the cylinder $h\Delta A$. And on the assumption that Eq. (2–11) is valid, we obtain

$$\begin{aligned} \mathbf{n} \cdot (\kappa_2\mathbf{E}_2 - \kappa_1\mathbf{E}_1) &= \frac{\sigma}{\epsilon_0}, \\ \mathbf{n} \cdot (\kappa_2\nabla\phi_2 - \kappa_1\nabla\phi_1) &= -\frac{\sigma}{\epsilon_0}. \end{aligned} \tag{2–15}$$

We have assumed that ΔA is small enough so that the fields are essentially constant.

The behavior of the tangential components of the fields as they cross the interface can be determined from the consideration of a small loop, as in Fig. 2–2, its major extent lying parallel to the surface, one side in medium 1 and the other in medium 2. From Eq. (1–34), we have for the closed line integral of the electric field

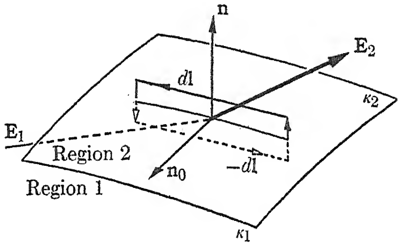


FIG. 2–2 Loop considered in determining the boundary conditions on the tangential components of **E**.

$$\oint \mathbf{E} \cdot d\mathbf{l} = 0. \tag{1–34}$$

If we apply this to the path indicated in the figure and let the ends of the loop shrink to zero, we have

$$\mathbf{E}_2 \cdot d\mathbf{l} - \mathbf{E}_1 \cdot d\mathbf{l} = 0. \tag{2–16}$$

In Fig. 2–2, **n** is again the unit normal to the interface and $\mathbf{n}_0$ is a unit vector normal to the loop of integration and lying in the surface. Since $d\mathbf{l}/dl = \mathbf{n}_0 \times \mathbf{n}$, and, with this substitution for $d\mathbf{l}$, Eq. (2–16) is a scalar triple product,

$$\mathbf{n}_0 \cdot \mathbf{n} \times (\mathbf{E}_2 - \mathbf{E}_1) = 0. \tag{2–17}$$

Now $\mathbf{n}_0$ may be oriented in any direction relative to **E**, and Eq. (2–17) holds for all orientations, so that

$$\mathbf{n} \times (\mathbf{E}_2 - \mathbf{E}_1) = 0, \tag{2–18}$$

and the tangential component of **E** is the same on both sides of the boundary. In terms of the potential,

$$\mathbf{n} \times (\nabla\phi_2 - \nabla\phi_1) = 0. \tag{2–19}$$

We have assumed that the loop is sufficiently short so that the fields are essentially constant over its length.

The relation (2–18) follows directly from Eq. (1–34) and an application of Stokes' theorem, but condition (1–34) is not necessary for the validity of Eq. (2–18). We shall see that the tangential component of **E** is continuous across a boundary even in the nonstatic case where curl **E** is not zero. In general, the boundary conditions on the normal component of **D** and the tangential component of **E** are implicit in the field equations, but the constitutive equations are necessary for deriving Eq. (2–15) and the relation for the tangential component of **D**. Since the tangential component of $\nabla\phi$ is continuous, we may put $\phi_1 = \phi_2$ at the boundary; the fact that the normal component of $\kappa\nabla\phi$ is continuous in the absence of a surface charge singularity completes the set of boundary conditions on the potentials and the fields on the two sides of an interface between two media.

2-3 The electric field in a material medium. We first defined the electrostatic field produced by free charges in a vacuum, and then we introduced material media containing charges that are inaccessible to measurement. The behavior of these media has been described in terms of their dipole moment per unit volume. Certain difficulties arise in the definition of the electric field within material media if one attempts to maintain a strictly phenomenological point of view. A definition of the field might be made by any of the following three methods, which will not necessarily be in agreement with each other:

(A) We may define the field on an atomic electron scale, where the question of the polarizability of material media would presumably not arise. Then for our macroscopic definition of the field we would take the space-time average of these atomic fields. A fast electron would experience such an average field.

(B) We might consider a hole cut in the dielectric material and define the field as that measured in this hole in terms of a unit charge such as was used in the vacuum definition. This cavity definition of the field will make the field strength depend on the geometry of the cavity and on its orientation with respect to the direction of the field in the medium. This will lead to a unique definition only if the shape and orientation of the cavity are standardized in an arbitrary way.

(C) We may define the field as that acting on an individual molecule of the dielectric.

Let us examine these methods separately.

(A) Space-time average definition. Consider a function $f(x,y,z;t)$ defined in a certain region of space during a certain time interval, as indicated in Fig. 2-3. The space-time average of $f(x,y,z;t)$ over a time interval $2T$ and a region of space of radius a is given by

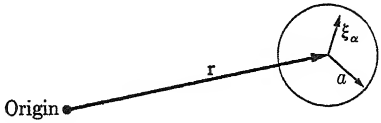


FIG. 2-3 Coordinates for averaging atomic fields. $\xi_\alpha = (\xi, \eta, \zeta)$.

$$\overline{f(x,y,z;t)} = \frac{1}{2T\frac{4}{3}\pi a^3}\int_{-T}^{T}\underbrace{\iiint}_{(\xi^2+\eta^2+\zeta^2)\leq a^2} f[(x+\xi),(y+\eta),(z+\zeta);(t+\theta)] \times d\xi\, d\eta\, d\zeta\, d\theta. \quad (2\text{-}20)$$

Performing this integral is a linear operation and may therefore be commuted with linear differential operators, as, for example,

$$\overline{\nabla f} = \nabla \bar{f}. \tag{2-21}$$

On an atomic scale an equation corresponding to Eq. (2-5) holds:

$$\nabla \cdot \mathcal{E} = \frac{\rho_a}{\epsilon_0}, \tag{2-22}$$

where $\mathcal{E}$ is the atomic electric field and ρ_a is the charge density in the atomic distribution. On taking the space-time average of ρ_a, we obtain

$$\overline{\nabla \cdot \mathcal{E}} = \frac{\bar{\rho}_a}{\epsilon_0} = \frac{\rho_t}{\epsilon_0}, \tag{2-23}$$

and from Eqs. (2-21) and (2-5),

$$\nabla \cdot \bar{\mathcal{E}} = \nabla \cdot \mathbf{E}. \tag{2-24}$$

Hence the macroscopic field **E** is actually the space-time average of the atomic field $\mathcal{E}$, even in the presence of dielectrics.

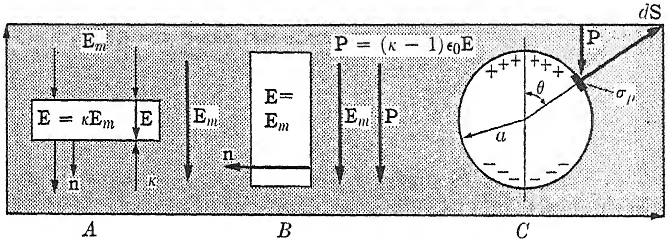

$\mathbf{E}_m$
$\mathbf{E} = \kappa\mathbf{E}_m$
E
n
κ
$\mathbf{E}_m$
$\mathbf{E} = \mathbf{E}_m$
n
$\mathbf{P} = (\kappa - 1)\epsilon_0\mathbf{E}$
$\mathbf{E}_m$
P
$d\mathbf{S}$
P
θ
σ_p
a
A
B
C

FIG. 2-4 Cavities for defining fields in a dielectric: slot *A* with short dimension parallel to the field; slot *B* with long dimension parallel to the field; and sphere *C*.

(B) Cavity definitions.

Consider the three shapes of cavities shown in Fig. 2-4. There are no true charges on the boundaries. From the boundary condition of Eq. (2-15) the field measured in slot *A*, whose major extent is oriented normal to the field, is $\kappa\mathbf{E}_m$, where $\mathbf{E}_m$ is the field in the medium. The field in slot *B*, whose major extent is oriented parallel to the field, is just $\mathbf{E}_m$, by Eq. (2-18). The field measured in the spherical cavity *C* can be shown to be

$$\mathbf{E} = \frac{3\kappa\mathbf{E}_m}{2\kappa + 1} = \mathbf{E}_m + \frac{\mathbf{P}}{\epsilon_0}\frac{1}{(2\kappa + 1)} \tag{2-25}$$

by methods to be discussed in Chapter 5 for the solution of boundary value problems. It is introduced here only to indicate how cavity defini-

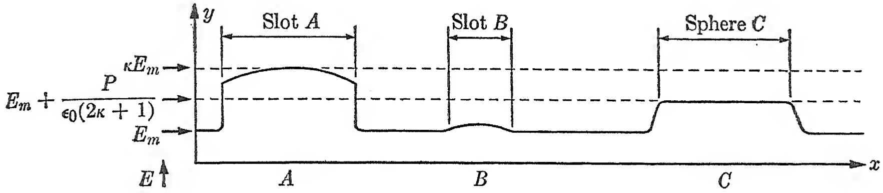


FIG. 2-5 Field profile measured by the cavity technique on a horizontal line passing through the centers of Fig. 2-4*A*, *B*, and *C*.

tions depend on the geometry of the cavity. For large values of κ Eq. (2-25) reduces to

$$\mathbf{E} = \tfrac{3}{2}\mathbf{E}_m. \tag{2-26}$$

The three types of fields existing in these cavities are shown graphically in Fig. 2-5. Each of the cavity definitions will thus give a definite value of the field, provided that the geometry is standardized.

(C) Molecular fields.

Consider a dielectric placed between the plates of a parallel plate condenser, as shown in Fig. 2-6, the dielectric and condenser being sufficiently large in the directions parallel to the plates so that end effects may be neglected. Consider one of the molecules constituting this dielectric. Let us draw a sphere of radius *a* about this particular molecule, intended to represent schematically the boundary between the microscopic and the macroscopic range of phenomena concerning the molecule. The molecule is thus influenced by the fields arising from the following charges:

(1) The charges on the surfaces of the condenser plates.
(2) The surface charge on the dielectric facing the condenser plates.
(3) The surface charge on the interior of the spherical boundary of radius *a*.
(4) The charges of the individual molecules, other than the molecule under consideration, contained within the sphere of radius *a*.

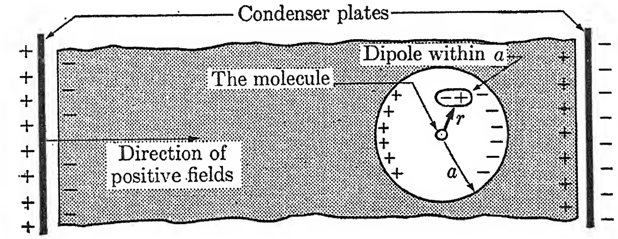


FIG. 2-6 Indicating the contributions of a dielectric to the field on one of its molecules.

The fields due to these sources may be computed separately:

(1) The charge on the condenser plates produces a field at the molecule in question equal to

$$\frac{\mathbf{D}}{\epsilon_0} = \mathbf{E} + \frac{\mathbf{P}}{\epsilon_0}. \tag{2-27}$$

(2) The polarization charge on the surface of the dielectric facing the condenser plates, $\sigma_P = P_n$, produces a field at the molecule given by

$$-\frac{\mathbf{P}}{\epsilon_0}. \tag{2-28}$$

(3) The polarization charge present on the inside of the sphere produces a field that may be calculated as follows. The magnitude of the electric field at the center of the cavity, $\mathbf{E}_P$, due to the polarization on the surface of the cavity, is given by

$$E_P = \frac{1}{4\pi\epsilon_0}\int \frac{\sigma_P \cos\theta \, dS}{a^2}, \tag{2-29}$$

where θ is the angle between $\mathbf{P}$ and the radius vector from the molecule to the surface element dS. The differential element of surface charge between θ and $\theta + d\theta$ is

$$\sigma_P \, dS = |\mathbf{P}| \, 2\pi a^2 \sin\theta \cos\theta \, d\theta. \tag{2-30}$$

Equation (2-29) thus becomes

$$E_P = \frac{P 2\pi a^2}{4\pi\epsilon_0 a^2}\int_0^{\pi} \sin\theta \cos^2\theta \, d\theta, \tag{2-31}$$

and on integrating, we obtain $E_P = P/3\epsilon_0$, or

$$\mathbf{E}_P = \frac{\mathbf{P}}{3\epsilon_0}. \tag{2-32}$$

Note that $\mathbf{E}_P$ is not the solution of the boundary value problem of a spherical cavity within a dielectric, as was the field given by Eq. (2-25), but is the solution of the problem of a spherical cavity within a dielectric if the polarization is considered to be unaffected by the presence of the cavity.

(4) The field due to the individual molecules within the sphere must be obtained by summing over the fields due to the dipoles within the sphere. We have the potential of an individual dipole from Chapter 1:

$$\phi = \frac{1}{4\pi\epsilon_0}\frac{\mathbf{p}\cdot\mathbf{r}}{r^3}. \tag{1-42}$$

The field at a distance r from a dipole is

$$\mathbf{E} = -\nabla\phi = \frac{-1}{4\pi\epsilon_0}\left[\frac{\mathbf{p}}{r^3} - \frac{3(\mathbf{p}\cdot\mathbf{r})\mathbf{r}}{r^5}\right]. \tag{2–33}$$

Summing over all the dipoles within the sphere is equivalent to taking the spatial average of the x-component of the field:

$$\overline{E_x} = \frac{-1}{4\pi\epsilon_0}\overline{\Sigma\left[\frac{p_x}{r^3} - \frac{3(p_x x^2 - p_y xy - p_z xz)}{r^5}\right]}. \tag{2–34}$$

Since according to our assumption the dielectric is isotropic, the x-, y-, and z-directions are equivalent, and

$$\overline{x^2} = \overline{y^2} = \overline{z^2} = \overline{r^2}/3, \quad \overline{xy} = \overline{yz} = \overline{zx} = 0. \tag{2–35}$$

Hence the field due to the dipoles within the sphere vanishes.

On adding the partial fields of expressions (2–27), (2–28), (2–32), and the zero field of Eq. (2–34), we have for the total field acting on one molecule

$$\mathbf{E}_{\text{eff}} = \mathbf{E} + \frac{\mathbf{P}}{\epsilon_0} - \frac{\mathbf{P}}{\epsilon_0} + \frac{\mathbf{P}}{3\epsilon_0} = \mathbf{E} + \frac{\mathbf{P}}{3\epsilon_0}. \tag{2–36}$$

This expression, derived for isotropic substances, is also valid for a lattice point within a cubic crystal, but is not valid for crystals of lower symmetry. Note that we have considered only dipole-dipole interactions between neighbors. Clearly, this will be inaccurate for substances having large oriented molecular groups.

The difference between Eq. (2–36) and method (A) is that here we consider what happens to an actual molecule of the medium, rather than take an average of the field at a random point. The physical significance of the space-time average of *all* the atomic fields would be the average field on a fast moving charge traversing the medium.

2–4 Polarizability. The field definition (C) is useful for describing the large-scale behavior of a dielectric in terms of the constants of its molecules. In order that such a description be made, the specific inductive capacity must be associated with the polarizability of a single molecule. This connection may be made by means of Eq. (2–36), which gives the field and thus the force acting on a single molecule within the body of a dielectric in terms of the external field. The quantity α, called the polarizability, is defined by the equation

$$\mathbf{p} = \alpha\epsilon_0\mathbf{E}_{\text{eff}}, \tag{2–37}$$

where $\mathbf{p}$ is the dipole moment induced in a single molecule. If N is the number of molecules per unit volume, the total polarization is

$$\mathbf{P} = N\mathbf{p} = \alpha N \epsilon_0 \mathbf{E}_{\text{eff}}$$
$$\mathbf{P} = \frac{N_0 g \alpha}{M} \epsilon_0 \mathbf{E}_{\text{eff}}, \tag{2-38}$$

where g is the density and M the molecular weight of the material, and N_0 is Avogadro's number, so that $N = N_0 g/M$. Therefore $\mathbf{P}$ may be found if α is given for a particular material. Furthermore, by combining Eq. (2–38) with Eqs. (2–12) and (2–36), we may eliminate the fields and obtain the relation sought:

$$\frac{\kappa - 1}{\kappa + 2} = \frac{\alpha N}{3} = \frac{N_0 g \alpha}{3M}. \tag{2-39}$$

This formula, known as the Clausius-Mossotti relation, gives the correct dependence of the specific inductive capacity on density for a wide class of solids and liquids. For dilute gases, where κ is not very different from unity, Eq. (2–39) becomes

$$\kappa - 1 = N\alpha = \frac{N_0 g \alpha}{M}, \tag{2-40}$$

just as would be expected for an approximation corresponding to the neglect of the interaction between each molecule and its neighbors. The molecular polarizability in general arises from two basic physical causes: (1) the lengthening of the bonds between atoms, and (2) the preferred orientation of molecules along the direction of the field as opposed to the random orientations brought about by thermal motions.

It is this second effect which is responsible for the temperature dependence of the specific inductive capacity. In statistical mechanics it is shown that under conditions of thermal equilibrium the probability that any one molecule has energy U is proportional to $e^{-U/kT}$, where k is Boltzmann's constant and T is the absolute temperature. If we have a molecule of intrinsic moment $\mathbf{p}_0$ in a field $\mathbf{E}$, then according to Section 1–8,

$$U = -\mathbf{p}_0 \cdot \mathbf{E}$$
$$= -p_0 E \cos\theta,$$

which depends on the orientation of the molecule with respect to the field. The contribution of each molecule to the total dipole moment would be $p_0 \cos\theta$, to be summed over all molecules. If there are N molecules per unit volume, the mean effective polarization would be given by

$$P = \frac{\displaystyle\int N p_0 \cos\theta e^{x\cos\theta}\, d\Omega}{\displaystyle\int e^{x\cos\theta}\, d\Omega}, \tag{2–41}$$

where for convenience we have written $x = p_0E/kT$. The denominator is easily integrated, and apart from the constant factor Np_0 the numerator is just the partial derivative of the denominator with respect to x. The final evaluation gives

$$P = Np_0\left(\coth x - \frac{1}{x}\right), \tag{2–42}$$

which for small x becomes

$$P \simeq Np_0 \frac{x}{3} = \frac{Np_0^2E}{3kT}. \tag{2–43}$$

This effect is to be added to the polarizability due to distortion of the molecules by the field, which we may call α_0. Equation (2–40) thus becomes

$$\kappa - 1 = N\alpha_0 + \frac{Np_0^2}{3\epsilon_0 kT}. \tag{2–44}$$

In the chapters immediately following we shall return to the problem of the electrostatic field in vacuum, but we are now able to consider dielectrics both as boundaries and as sources.

SUGGESTED REFERENCES

All the references listed at the end of Chapter 1 for the fundamental principles of vacuum electrodynamics contain treatments of dielectrics and boundary conditions. For the derivation of Eq. (2–42), called the Langevin formula, see almost any textbook on statistical mechanics, for example:

R. H. Fowler, *Statistical Mechanics.* Uses the elegant methods involving the partition function.

J. E. Mayer and M. G. Mayer, *Statistical Mechanics.* Mathematically simpler than Fowler, and more than adequate for our purpose.

EXERCISES

1. The fact that the magnitude of **P** is proportional to that of **E**, and that the angle between **P** and **E** is constant for a given orientation of a crystal dielectric in the field, can be expressed by

$$\mathbf{P} = (\mathbf{e}\cdot\mathbf{e})\mathbf{E} - (\mathbf{e}\cdot\mathbf{E})\mathbf{e}.$$

Find **e** in terms of **P**, **E**, and the angle between **P** and **E**, and show that $\chi_{\alpha\beta}$ in Eq. (2–13), $P_\alpha = \epsilon_0 \chi_{\alpha\beta} E_\beta$, is a symmetric tensor.

2. Consider a simple cubic lattice of point dipoles, all of equal moment and like orientation. Show that the field at the position of one dipole due to all others in a sphere of arbitrary radius about this point is zero.

3. A long very thin rod of dielectric constant κ is oriented parallel to a uniform external field **E**. What are **E** and **D** within the rod? What are the fields in a very thin disk of dielectric oriented perpendicular to the field?

4. Show that for an electret (fixed **P**) the integral $\int \mathbf{E} \cdot \mathbf{D}\, dv$ over the entire field volume vanishes.

5. Consider an electron of charge $-e$ moving in a circular orbit of radius a_0 about a charge $+e$ in a field directed at right angles to the plane of the orbit. Show that the polarizability α is approximately $4\pi a_0^3$.

CHAPTER 3

GENERAL METHODS FOR THE SOLUTION OF POTENTIAL PROBLEMS

We have seen that, in principle, potential problems are solvable if all charge distributions are known, and we shall prove the uniqueness of any solution which reduces to the correct values on the boundary of a region. Actually only a few of the most idealized problems can be solved with any degree of simplicity. For practical applications experimental and numerical methods of mapping fields have been devised, as well as graphical and semigraphical procedures involving some calculations. Analytic methods rarely lead to solutions in closed form, while infinite series must converge fairly rapidly to be useful. Nevertheless, analytic solutions of geometrical boundary configurations approximating the actual situation furnish a valuable check even when experimental mapping is finally necessary, and graphical methods depend on prior knowledge of the general behavior of the potential. The solution of problems with relatively simple boundaries and charge distributions is therefore of value, even for more complicated engineering applications.

Unfortunately, no general methods of solution are available which will apply to all types of geometrically simple problems, and therefore each individual case demands, to some extent, special treatment. Certain methods apply to general classes of problems, however, and can be discussed as individually characteristic of these classes.

3–1 Uniqueness theorem. This theorem states that if within a given boundary a solution of a potential problem is found which reduces to the given potential distribution on that boundary, or to the given normal derivative of the potential on that boundary, then this solution is the only correct solution of the potential equations within the boundary. The theorem provides justification for attempting any method of solution so long as the resulting solution can be shown to obey Laplace's equation in a charge-free region. No matter how the solution is obtained, if it satisfies these conditions the problem is considered solved.

The proof of the theorem is very similar to that indicated (Section 1–1) for the unique definition of vector fields from their source and circulation densities, except that here the integration does not extend to infinity. If we put $\phi\nabla\phi$ into Gauss's divergence theorem as the vector field, we obtain

$$\int \phi\nabla\phi \cdot d\mathbf{S} = \int \nabla \cdot (\phi\nabla\phi)\, dv = \int [(\nabla\phi)^2 + \phi\nabla^2\phi]\, dv. \qquad (3\text{–}1)$$

The last term vanishes, from Laplace's equation, if we choose the surface of integration in such a way as to exclude all charged regions from the interior of the region of integration. It may be necessary to employ surfaces internal to the outer boundary in order to exclude the charges from v.

Let us suppose that two different potentials, ϕ_1 and ϕ_2, are solutions of a given potential problem. Both ϕ_1 and ϕ_2 are to satisfy the boundary condition, and hence on the boundary either $\phi_1 = \phi_2$ or $(\nabla\phi_1)_n = (\nabla\phi_2)_n$. (The component of $\nabla\phi$ normal to a surface is often called the "normal derivative" and may be designated by $\partial\phi/\partial n$.) If we substitute the difference $\phi_1 - \phi_2$ for ϕ in Eq. (3–1), we have

$$\int (\phi_1 - \phi_2)\nabla(\phi_1 - \phi_2) \cdot d\mathbf{S} = \int [\nabla(\phi_1 - \phi_2)]^2 \, dv. \tag{3–2}$$

Either boundary condition (equality of the potentials or their normal derivatives) assures the vanishing of the left side of Eq. (3–2). Since the integrand of the right side of Eq. (3–2) is positive definite, it must be zero in order for the integral to vanish; hence throughout the volume v

$$\nabla\phi_1 = \nabla\phi_2, \quad \phi_1 = \phi_2 + \text{constant}. \tag{3–3}$$

Thus the two potentials that were assumed to be different yet satisfying the same boundary condition can differ at most by an additive constant which makes no contribution to the gradient; therefore these potentials will give the same electric field distributions.

If linear dielectrics are involved, Eq. (3–1) may be replaced by

$$\int \phi \, \kappa\nabla\phi \cdot d\mathbf{S} = \int [\kappa(\nabla\phi)^2 + \phi\nabla \cdot (\kappa\nabla\phi)] \, dv. \tag{3–1′}$$

Laplace's equation for dielectrics is

$$\nabla \cdot (\kappa\nabla\phi) = 0,$$

and hence the proof for uniqueness remains valid. If non-linear dielectrics are involved, the region may be divided into subregions having uniform polarization densities, and for which the theorem holds separately.

3–2 Green's reciprocation theorem. A large number of theorems that are useful for the solution of electrostatic problems serve to transform the solution of a known, presumably simpler, problem into the solution of another problem whose solution is desired. Of such theorems one of the most useful is Green's reciprocity theorem. Let us consider a set of n point charges q_j, at positions where the potentials due to the other charges are given by a set of numbers ϕ_j. The potential at the point j is related to the charges at the other points (designated by i for purposes of summation) by

$$\phi_j = \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{n}{}' \frac{q_i}{r_{ij}}. \tag{3-4}$$

The prime on the summation sign means that the term $i = j$ is to be omitted from the summation. If, on the other hand, a different set of charges q_j' is placed at the same points, giving rise to the corresponding potentials ϕ_j', a similar relation holds:

$$\phi_j' = \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{n}{}' \frac{q_i'}{r_{ij}}. \tag{3-5}$$

Let us now multiply Eq. (3-4) by q_j' and Eq. (3-5) by q_j, then sum each expression over the index j:

$$\begin{aligned} \sum_{j=1}^{n} \phi_j q_j' &= \sum_{j=1}^{n} \sum_{i=1}^{n}{}' \frac{q_i q_j'}{r_{ij}} \frac{1}{4\pi\epsilon_0}, \\ \sum_{j=1}^{n} \phi_j' q_j &= \sum_{j=1}^{n} \sum_{i=1}^{n}{}' \frac{q_i' q_j}{r_{ij}} \frac{1}{4\pi\epsilon_0}. \end{aligned} \tag{3-6}$$

Since i and j are summation indices, we may interchange them in one product of the q's, and see that

$$\sum_{j=1}^{n} \phi_j q_j' = \sum_{j=1}^{n} \phi_j' q_j, \tag{3-7}$$

which is the desired theorem.

This theorem can be generalized from a set of point charges to a set of n conductors of potentials ϕ_j carrying charges q_j: the generalization follows if we combine the points of equal ϕ_j in Eq. (3-7) into a single term. Equation (3-7) thus applies directly to such a system of conductors. If all but two conductors i and j are grounded, Eq. (3-7) implies that the potential to which the uncharged conductor i is raised by putting a charge q on conductor j is equal to the potential of j, when uncharged, produced by a charge q on i. An application of Eq. (3-7) to the solution of a potential problem is indicated at the close of Section 3-3.

3-3 Solution by Green's function. A great variety of solutions of potential problems can be generated from what is known as a Green's function. The Green's function for a particular geometrical arrangement is the solution of the potential problem for this given geometrical arrangement of grounded conducting boundaries when the only charge present is a unit point charge at point x_α. It should be noted that the grounded conducting boundaries may be at infinity and the point charge need not be surrounded by a zero potential surface at a finite distance. It may be shown with the aid of Green's reciprocity theorem that the Green's function for a particular geometry is a symmetrical function of the coordinates of a unit charge located at the point x'_α and the coordinates of the point of observation x_α.

Two general types of problems can be solved by the use of Green's function. One type is that in which the potential distribution over a certain conducting boundary is given, and the other is one in which the charge distribution is given in a region within a conducting boundary. The derivation of the solution of both these problems can be given together by means of Green's theorem,

$$\int (\phi\nabla^2\psi - \psi\nabla^2\phi)\,dv = \int (\phi\nabla\psi - \psi\nabla\phi)\cdot d\mathbf{S}, \tag{3–8}$$

where ψ and ϕ are arbitrary functions of position which are required to be nonsingular throughout the volume v. Let ϕ be the desired solution of a particular potential problem and let $\psi = G$ be the Green's function for the geometry of the problem, i.e., the solution of the problem of a unit point charge located at $r = 0$ with the surface S grounded. Then G will be of the form

$$G = \frac{1}{4\pi\epsilon_0 r} + \chi, \tag{3–9}$$

where χ represents the potential due to the induced charge on S. Here χ is harmonic in v, i.e., it is a solution of Laplace's equation. Therefore, G has a singularity only at $r = 0$, which we may handle by means of the δ-function. On substituting Eq. (3–9) into Green's theorem, we have

$$\int (G\nabla^2\phi - \phi\nabla^2 G)\,dv = \int \left[G\nabla^2\phi + \frac{\phi\delta(r)}{\epsilon_0}\right] dv$$

$$= \int (G\nabla\phi - \phi\nabla G)\cdot d\mathbf{S}. \tag{3–10}$$

Also, by definition, $G = 0$ on S. Hence, on collecting the nonvanishing terms, we find

$$\phi = -\epsilon_0\left(\int G\nabla^2\phi\,dv + \int \phi_s\nabla G\cdot d\mathbf{S}\right). \tag{3–11}$$

Let us now consider the two cases mentioned earlier:

(1) The surface surrounding the point x'_α is grounded, making $\phi_s = 0$, and $\nabla^2\phi = -\rho/\epsilon_0$ due to the charge distribution ρ throughout v. Equation (3–11) then reduces to

$$\phi(x'_\alpha) = -\epsilon_0\int G\nabla^2\phi\,dv = \int G\rho\,dv. \tag{3–12}$$

This expression is fairly obvious, since it merely represents the principle of superposition applied to the density of point sources within the volume v, with each unit source of which the density ρ consists contributing its share to the potential $\phi(x'_\alpha)$ by the superposition indicated by the integral.

(2) Let there be no sources of ϕ throughout the volume v, so that $\nabla^2\phi = 0$, but let us assume that ϕ is a given function ϕ_s on the surface S. In this case, Eq. (3–11) reduces to

$$\phi(x'_\alpha) = -\epsilon_0 \int \phi_s \nabla G \cdot d\mathbf{S}. \tag{3–13}$$

Equation (3–13) gives the potential within a given region enclosed by a conducting boundary where different parts of the boundary are raised to a given set of potentials. This solution expresses the potential within this boundary in terms of the surface integral of the potential on the boundary multiplied by the normal derivative of the Green's function. Physically, the normal derivative of the Green's function represents the surface charge density that is induced on the grounded conducting boundary by a unit charge at the point x'_α. Equation (3–13) thus gives the solution of the potential problem corresponding to a given potential on the boundary in terms of the integral of this potential multiplied by the charge induced on the grounded boundary by a unit charge placed at the field point. If we wish to express Eq. (3–13) explicitly in terms of the charge σ_{1s} induced on the grounded boundary we note from Eq. (2–15) that

$$\nabla G \cdot \frac{d\mathbf{S}}{dS} = +\frac{\sigma_{1s}}{\epsilon_0}. \tag{3–14}$$

Thus Eq. (3–13) becomes simply

$$\phi(x'_\alpha) = -\int \phi_s \sigma_{1s}\, dS. \tag{3–15}$$

Theorem (3–15) may also be derived directly by the use of Green's reciprocation theorem:

(1) Let the surface S be grounded and let a single charge $q_{x'_\alpha}$ be located at the point x'_α. The charge induced by $q_{x'_\alpha}$ on the jth region of the boundary S will be designated by q_{1_j}.

(2) Let the charge at x'_α be removed, but let the surface S be divided into sections, each at a constant potential, the potential of the jth section of S being ϕ_{js}. In this case $\phi(x'_\alpha)$ represents the potential at x'_α.

If we relate these two cases by means of Eq. (3–7), we obtain

$$q_{x'_\alpha}\phi(x'_\alpha) + \sum_{j=1}^{n} q_{1j}\phi_{js} = 0 + 0. \tag{3–16}$$

The two zeros on the right arise from the fact that the potential is zero over the entire boundary in case (1), and that the charge at x'_α is zero in case (2). Remembering that $q_{x'_\alpha}$ is unity in our consideration, we have

$$\phi(x'_\alpha) = -\sum_{j=1}^{n} \phi_{js} q_{1j}. \tag{3-17}$$

This expression is identical with Eq. (3–15), but it has been obtained directly in terms of the induced charge in a way that is more obvious physically.

In the consideration of more specific problems, we shall derive Green's functions for various conducting boundaries. The solutions for the problems, both for grounded conductors enclosing charge distributions and for charge-free regions surrounded by conductors whose potentials are given, can then be written down immediately.

3–4 Solution by inversion. There are various kinds of transformations by which a set of solutions of one potential problem can be transformed into the solutions of another problem. The process of inversion is a special case, important because it is valid in three dimensions as well as in two dimensions. In two dimensions classes of such transformations more general than inversion can be found.

One of the simplest and most useful methods by which the solution of a problem can often be transformed into the solution of a simpler problem is the inversion in a sphere, as shown in Fig. 3–1. It can be shown by direct differentiation that if $\phi_p = \phi(r,\theta,\varphi)$ is a solution of Laplace's equation, then $\psi_p = \psi(r',\theta,\varphi) = \frac{a}{r'}\,\phi\left(\frac{a^2}{r'},\theta,\varphi\right)$ is also a solution of Laplace's equation. In relation to a sphere of radius a this transformation of the point r into the point r', by the relation $rr' = a^2$, maps the point $p(r,\theta,\varphi)$ into its inversion point $p'(a^2/r,\theta,\varphi)$, moving the point along the radius vector from a position inside the sphere to a point outside, or vice versa. Let a charge

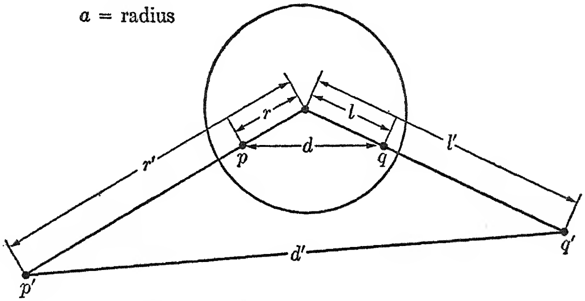


FIG. 3–1 Solution by inversion.

q be placed at distance l and a charge q' at distance l' from the center such that $ll' = a^2$. The relations $rr' = a^2$ and $ll' = a^2$ imply that $r/l' = l/r'$, and therefore the triangles with sides rld and $r'l'd'$ are similar. Thus we have $r/l' = l/r' = d/d'$. The potential at p before inversion is $\phi_p = q/4\pi\epsilon_0 d$ and the potential at p' after inversion is $\phi_{p'} = q'/4\pi\epsilon_0 d'$, so that

$$\frac{\phi_p}{\phi_{p'}} = \frac{q'}{d'}\frac{d}{q} = \frac{q'}{q}\frac{l}{r'} = \frac{q'}{q}\frac{r}{l'}. \tag{3–18}$$

To formulate a law for the inversion of charges, we make use of the fact that zero potential surfaces must transform into zero potential surfaces. We can make the potential of the inversion sphere zero by taking two charges initially, $q_{\text{I}} = q$ at l, and $q_{\text{II}} = -qa/l$ at a distance l' from the center such that $ll' = a^2$. Now the inversion sphere at zero potential under the influence of the two charges is to remain so after inversion. This is assured if the two charges change places thus:

$$q_{\text{I}} = q \text{ at } l \text{ becomes } q'_{\text{I}} = aq/l \text{ at } a^2/l,$$

$$q_{\text{II}} = -qa/l \text{ at } a^2/l \text{ becomes } q'_{\text{II}} = -q \text{ at } l.$$

In either case the transformed charge is the original charge multiplied by the inversion radius a over the original distance of the charge from the center of the sphere,

$$\frac{q'}{q} = \frac{a}{l} = \frac{l'}{a} = \sqrt{\frac{l'}{l}}. \tag{3–19}$$

It seems more convenient for a charge to retain its original sign and change only its magnitude when it is inverted, although this is not necessary if all charges undergoing inversion are treated in the same way. We now secure, by substituting Eq. (3–19) into Eq. (3–18), the rule for the inversion of potentials,

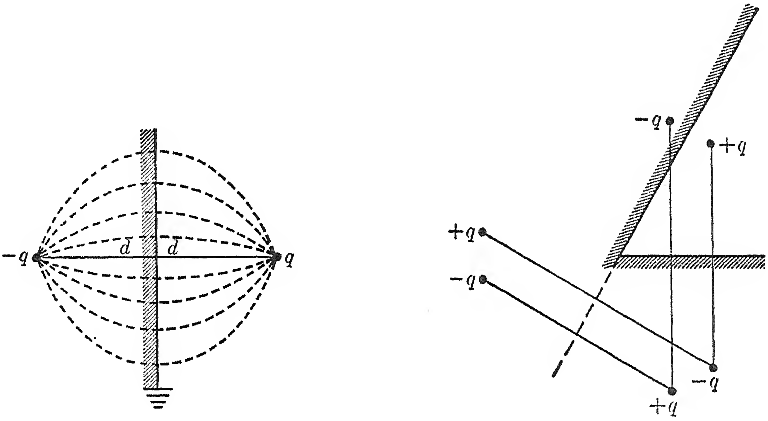
$$\phi_{p'}/\phi_p = a/r' = r/a, \tag{3–20}$$

in agreement with $\phi_{p'} = \dfrac{a}{r'}\,\phi(r,\theta,\varphi)$. The transformation equations for such quantities as volume or surface charge densities can be obtained by multiplying the charge transformation Eq. (3–19) by the transformation of the appropriate geometrical quantities.

In an inversion transformation a point charge will often appear at the center of inversion in the transformed geometry. This point charge arises from the fact that the net charge in the original geometry had electric field lines that terminated on equal and opposite charges located at infinity and, in the inversion, infinity is brought in to the origin.

The main utility of the inversion transformation is that it rectifies spherical boundaries if the center of inversion is taken on the spherical boundary. Two freely charged intersecting spheres may be inverted into two intersecting planes, and the plane boundary problem is usually easily soluble by the method of images.

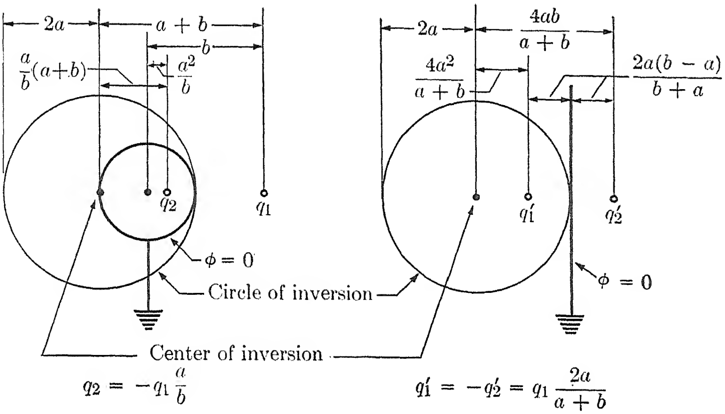
3–5 Solution by electrical images. If we have two equal and opposite point charges, the zero potential surface is the plane of points equally distant from the two charges. The zero potential plane could be replaced by a plane grounded conductor, and the potential and field on the left in Fig. 3–2(a) would remain unchanged. Thus the solution for a point charge and grounded plane is just that of two point charges throughout the region of space in which the field exists. The fictional charge $-q$ is called the "image" of q, by analogy with reflection in a mirror. A case in which two semi-infinite intersecting plane conductors produce five images is shown in Fig. 3–2(b).



(a) (b)

FIG. 3–2 Examples of charge "images" in plane grounded conductors.

The method of inversion discussed above justifies the solution of the problem of a point charge opposite a grounded conducting sphere, as shown in Fig. 3–3(a), by the method of electrical images. The sum of the potential from a point charge q_1 located in free space near a grounded sphere of radius a, and from the charge that is induced on this sphere by q_1, will be a potential distribution in which the sphere is a zero potential surface. This system can be transformed into q_1' and a plane by inversion in the sphere of radius $2a$ indicated in the figure. The uniqueness theorem requires that the potential outside the original spherical conductor will be identical with that of the reciprocally transformed point charge and plane, or equivalent point charge and image. But the plane conductor image transforms into q_2, which may be called the image of q_1



(a) (b)

FIG. 3-3 Inversion of a point charge and conducting sphere, (a), into a point charge and conducting plane, (b).

in the conducting sphere. The potential corresponding to the point charge and the image it makes in the grounded sphere is therefore the correct Green's function in the region bounded by the conducting sphere and by infinity.

By combination of the methods of electrical images and inversion one can readily write down the solutions to problems of the type in which various areas of the surface of the conducting sphere are raised to different arbitrary potentials, since the problems involving a plane of which various areas are raised to different potentials can be easily solved by using the potential of the point charge and its image as a Green's function. Problems in which there is an arbitrary charge distribution in the region outside of a grounded conducting sphere are also immediately transformable to plane problems and solved by the image method. Similarly corresponding solutions to problems involving the inside of a grounded conducting sphere are obtained. Even problems involving two or three intersecting spheres can be transformed to problems of intersecting planes if the center of inversion is taken at a point of intersection of the spheres.

3-6 Solution of Laplace's equation by the separation of variables. Except within a distribution of charge, the fundamental problem of potential theory is to find a solution of Laplace's equation which satisfies certain conditions on the boundaries of the region under consideration. If these boundaries correspond to coordinate surfaces in a system of orthogonal coordinates, the solution by separation of variables is often much more convenient than the general Green's function method. For one thing, it is very easy to state the boundary conditions in the appropriate system of coordinates, whether the condition be continuity of ϕ or its derivative or the assignment of some definite value of ϕ or $\partial\phi/\partial n$, when each statement refers to a constant value of a particular coordinate. This would be true of any system of coordinates, but in certain systems we can go further and write the solution as a product of functions of the coordinates separately, so that the boundary conditions can be applied to the separate single-variable factors. It may be added that while there is no direct general method of solving partial differential equations the separation reduces Laplace's equation to a set of ordinary differential equations which in principle are always solvable.

The essential features of the method can best be demonstrated by means of an example. Consider a pair of parallel grounded conducting plates at $y = 0$ and $y = a$, as shown, with a line charge parallel to the z-axis at the point $(0,d)$. We seek a solution valid between the plates, assuming there are no other charges. The problem is thus a two-dimensional one for which rectangular coordinates are appropriate. Let us assume that $\phi(x,y) = X(x)Y(y)$, where X is a function of x alone and Y is a function of y. Except for the point $(0,d)$ the equation to be satisfied is

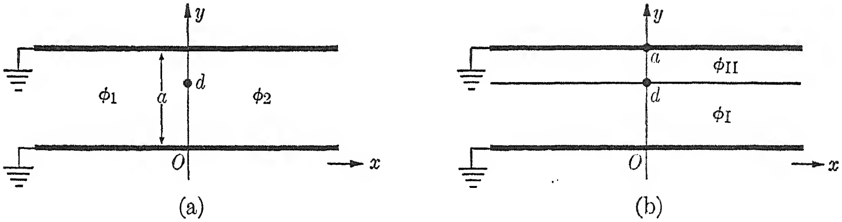


FIG. 3–4 Showing two ways of dividing the space between condenser plates by planes containing the line of charge at $x = 0$, $y = d$.

$$\nabla^2\phi = YX'' + XY'' = 0, \tag{3–21}$$

where the double prime denotes the second derivative of the function with respect to its argument. If we divide Eq. (3–21) by ϕ, we obtain

$$\frac{X''}{X} + \frac{Y''}{Y} = 0. \tag{3–22}$$

Since x and y can vary independently, both terms of Eq. (3–22) must be independent of either variable, and we can write

$$X''/X = -Y''/Y = C. \tag{3–23}$$

The constant C is called the separation parameter. If there are no restrictions on C the product of the general solutions of the ordinary differential equations is a general solution of the two-dimensional Laplace equation. The boundary conditions of the physical problem, however, will limit both the nature of the solutions and the values of the separation parameter. The solution we seek is a sum (or integral, depending on whether the allowed values of the parameter are discrete or continuous) of allowed product solutions, with coefficients determined so that the boundary conditions are exactly satisfied. In order to determine these coefficients we shall make use of a property of the functions known as *orthogonality*.

There remains a choice in the sign of the separation constant C and thus in the nature of the corresponding solutions. Let us first assume that C of Eq. (3–23) is positive, $C = k^2$, so that the ordinary differential equations become

$$\begin{aligned} Y'' + k^2 Y &= 0, \\ X'' - k^2 X &= 0, \end{aligned} \tag{3–24}$$

having general solutions

$$\begin{aligned} Y &= A \sin ky + B \cos ky, \\ X &= Ce^{kx} + De^{-kx}. \end{aligned} \tag{3–25}$$

The boundary conditions to be satisfied are that $\phi = 0$ at $y = 0$, $y = a$, and $x = \pm\infty$. The potential may be made to vanish at the plates simply by setting $B = 0$ and limiting k to the values $n\pi/a$, where n is an integer. The conditions at plus and minus infinity along x, however, cannot be simultaneously satisfied by either term of X, so that we must write the solutions separately for the regions of positive and negative x:

$$\begin{aligned} \underset{-\infty<x<0}{\phi_1} &= \sum_{n=1}^{\infty} C_n e^{n\pi x/a} \sin \frac{n\pi y}{a}, \\ \underset{0<x<\infty}{\phi_2} &= \sum_{n=1}^{\infty} A_n e^{-n\pi x/a} \sin \frac{n\pi y}{a}. \end{aligned} \tag{3–26}$$

At $x = 0$ ϕ is continuous, so that the coefficients in the two series are equal term by term, i.e., $C_n = A_n$. We have yet to determine these coefficients, however, and we have not taken account of the flux from the line charge. The physical requirement must exactly correspond to the mathematical determination of the A_n's, since the potential is unique.

Now a line charge could be represented by a two-dimensional δ-function, but since we are writing the solution separately for regions 1 and 2 it is possible to use our knowledge of the boundary conditions at a surface charge and employ a one-dimensional function $\delta(y - d)$ as a special case of an arbitrary charge distribution along the surface $x = 0$. In other words, on $x = 0$,

$$\sigma(y) = q\,\delta(y - d), \tag{3–27}$$

where q is the charge per unit length perpendicular to the xy-plane and $\delta(y-d)$ is defined by the equations

$$\int_0^a \delta(y-d)\,dy = 1,$$
$$\int_0^a f(y)\,\delta(y-d)\,dy = f(d), \quad 0 < d < a. \tag{3-28}$$

The potentials must then satisfy the conditions

$$\frac{\sigma(y)}{\epsilon_0} = \frac{q\,\delta(y-d)}{\epsilon_0} = \left[\frac{\partial\phi_1}{\partial x} - \frac{\partial\phi_2}{\partial x}\right]_{x=0} = \Sigma A_n \frac{2n\pi}{a} \sin\frac{n\pi y}{a}. \tag{3-29}$$

The Fourier coefficients are determined in the usual way by multiplying both sides of Eq. (3-29) by $\sin m\pi y/a$ and integrating from 0 to a. All terms of the series on the right will vanish except that for which $m = n$, i.e., the sine functions are "orthogonal" over the interval. Therefore

$$\frac{q}{\epsilon_0}\sin\frac{m\pi d}{a} = \frac{2m\pi}{a} A_m \frac{a}{2},$$

or, on writing n for m,

$$A_n = \frac{q}{\epsilon_0 n\pi}\sin\frac{n\pi d}{a}. \tag{3-30}$$

The entire solution is then

$$\phi_1 = \frac{q}{\epsilon_0\pi}\Sigma\frac{1}{n}\sin\frac{n\pi d}{a}\,e^{n\pi x/a}\sin\frac{n\pi y}{a},$$
$$\phi_2 = \frac{q}{\epsilon_0\pi}\Sigma\frac{1}{n}\sin\frac{n\pi d}{a}\,e^{-n\pi x/a}\sin\frac{n\pi y}{a}, \tag{3-31}$$

and the problem is solved.

It is instructive to note that the same potential may look quite different with the opposite choice of sign for the separation parameter in Eq. (3-23). If we put $C = -k^2$, the solution for $X(x)$ is just $\cos kx$, since the potential is obviously an even function of x, but no limitations are imposed on k. No single function Y will vanish on the two conducting plates, however, and we must again divide the region into two parts, this time by the plane $y = d$. It can be easily verified that for any k the two solutions which vanish at $y = 0$ and $y = a$, and are continuous at $y = d$, are

$$\begin{aligned} Y_{\mathrm{I}} &= \frac{\sinh ky}{\sinh kd}, \quad 0 < y < d, \\ Y_{\mathrm{II}} &= \frac{\sinh k(a-y)}{\sinh k(a-d)}, \quad d < y < a. \end{aligned} \tag{3-32}$$

The potentials are integrals over k, with coefficients which we may call $A(k)$:

$$\begin{aligned} \phi_{\mathrm{I}} &= \int_{-\infty}^{\infty} A(k) \cos kx \frac{\sinh ky}{\sinh kd}\, dk, \\ \phi_{\mathrm{II}} &= \int_{-\infty}^{\infty} A(k) \cos kx \frac{\sinh k(a-y)}{\sinh k(a-d)}\, dk. \end{aligned} \tag{3-33}$$

The charge density on the plane $y = d$ is now a function of x, and the condition on the potentials is

$$\begin{aligned} \frac{q\,\delta(x)}{\epsilon_0} &= \left[\frac{\partial \phi_{\mathrm{I}}}{\partial y} - \frac{\partial \phi_{\mathrm{II}}}{\partial y}\right]_{y=d} \\ &= \int A(k) \cos kx \left[\frac{\cosh kd}{\sinh kd} + \frac{\cosh k(a-d)}{\sinh k(a-d)}\right] k\,dk \\ &= \int A(k) \frac{\cos kx \sinh ka}{\sinh kd \sinh k(a-d)}\, k\,dk. \end{aligned} \tag{3-34}$$

But it follows from a formal application of the Fourier integral theorem, and has been indicated in a problem at the end of Chapter 1, that

$$\int_{-\infty}^{\infty} \cos kx\, dk = 2\pi\, \delta(x) \tag{3-35}$$

in the sense that both behave in the same way as factors in an integrand. Therefore

$$A(k) = \frac{q}{2\pi\epsilon_0} \frac{\sinh kd \sinh k(a-d)}{k \sinh ka} \tag{3-36}$$

and the potential is given explicitly by

$$\begin{aligned} \phi_{\mathrm{I}} &= \frac{q}{2\pi\epsilon_0} \int \frac{\sinh k(a-d)}{k \sinh ka} \cos kx \sinh ky\, dk, \\ \phi_{\mathrm{II}} &= \frac{q}{2\pi\epsilon_0} \int \frac{\sinh kd}{k \sinh ka} \cos kx \sinh k(a-y)\, dk. \end{aligned} \tag{3-37}$$

If $q = 1$ this solution [Eq. (3–31) or (3–37)] is the Green's function for the two-dimensional parallel plate geometry, and we shall see that the method is general for determining the Green's function for a set of equicoordinate planes. But for our immediate purpose the details of this particular problem are less important than the principles they illustrate. Usually one choice of parameters and corresponding functions is more convenient than the other, but a choice must be made.

To achieve separation of variables in three dimensions we follow the same procedure. In Cartesian coordinates $\phi(x,y,z) = X(x)Y(y)Z(z)$, and substitution in Laplace's equation yields

$$\frac{\nabla^2\phi}{\phi} = \frac{X''}{X} + \frac{Y''}{Y} + \frac{Z''}{Z} = 0. \tag{3–38}$$

We may therefore set

$$\begin{aligned} \frac{X''}{X} + \frac{Y''}{Y} &= -\frac{Z''}{Z} = C_1, \\ \frac{X''}{X} &= C_1 - \frac{Y''}{Y} = C_2. \end{aligned} \tag{3–39}$$

There are now two separation parameters; in general, the number of such parameters corresponding to N independent variables is $N - 1$. The lack of symmetry in the equations for the factor solutions is intrinsic in Laplace's equation, and may be said to correspond to the complete symmetry in sign of the coordinates themselves in the Laplacian operator. In three-dimensional rectangular coordinates two of the factors may be circular functions, but not all three. It is necessary that the functions be orthogonal in order to permit the determination of coefficients at a constant surface of the third variable, but this can be shown to be a general property of solutions of Laplace's equation, independent of the coordinate system for every case in which separation is possible.

The orthogonality of the allowed functions may be demonstrated by means of Green's theorem. Let us apply the proof in the case of spherical polar coordinates, where the geometry is easily visualized, and then see that it is applicable to other coordinate systems as well. Assume that Laplace's equation is separable, so that $\phi(r,\theta,\varphi) = R(r)Y(\theta,\varphi)$, and that whatever the nature of the functions or the parameter involved in the separation, ϕ_1 and ϕ_2 are two allowed solutions. If we put these two potentials into Green's theorem, we obtain

$$\int (\phi_1\nabla^2\phi_2 - \phi_2\nabla^2\phi_1)\,dv = \int [R_1Y_1\nabla(R_2Y_2) - R_2Y_2\nabla(R_1Y_1)]\cdot d\mathbf{S}. \tag{3–40}$$

The left side vanishes, since ϕ_1 and ϕ_2 are solutions of Laplace's equation. If we let S be the surface of a sphere, the component of ∇ parallel to S does not operate on the function Y, and therefore

$$\left(\frac{R_2'}{R_2} - \frac{R_1'}{R_1}\right)\int Y_1 Y_2\,dS = 0. \tag{3-41}$$

But if the two radial functions correspond to different values of the separation parameter they have different dependence on r, and thus their logarithmic derivatives are unequal. Equation (3-41) is thus a statement that the two functions Y_1 and Y_2 are orthogonal to each other when integrated over the surface of a sphere.

If the coordinate surface is not closed the proof follows in exactly the same way except that use must be made of the fact that the potentials are zero at infinity. For sources confined to a finite region the potentials do approach zero sufficiently fast so that the integral over the infinite surface vanishes. We can therefore conclude that, in general, orthogonal functions are generated in the solution of Laplace's equation, and that if the equation is separable it is possible in principle to complete the solution.

It can be shown that the set of allowed functions is complete as well as orthogonal. We shall not take up the proof of this statement, but it should be remarked that completeness is necessary for the existence of satisfactory potential theory. Without it the representation of arbitrary potential or charge distributions on surfaces could not be guaranteed.

SUGGESTED REFERENCES

E. Weber, *Electro-magnetic Fields*, Vol. 1, *Mapping of Fields*. This excellent treatise on practical solutions also presents the principles in a clear and coherent fashion.

O. D. Kellogg, *Foundations of Potential Theory*. The classical presentation of mathematical foundations.

J. C. Maxwell, *Electricity and Magnetism*, Vol. 1. Although originated by Thomson, the method of images and inversion was further developed by Maxwell, and his account is excellent.

Of the works listed at the end of Chapter 1 the most useful for electrostatics problems are those by Jeans, Smythe and Stratton.

P. Morse and H. Feshbach, *Methods of Mathematical Physics*. This recent and most comprehensive work on mathematical methods has almost made other mathematical references unnecessary.

E. T. Whittaker and G. N. Watson, *Modern Analysis*. Much more formal mathematically than Morse and Feshbach, this work is generally useful for the properties of various functions and especially here for a discussion of convergence.

Somewhat more elementary treatments of useful mathematical topics are:

W. E. Byerly, *Fourier's Series and Spherical Harmonics*.

R. V. Churchill, *Fourier Series and Boundary Value Problems*.

EXERCISES

1. If $\phi(x,y,z)$ is a solution of Laplace's equation, show that $\frac{1}{r}\phi\left(\frac{x}{r^2}, \frac{y}{r^2}, \frac{z}{r^2}\right)$ is also a solution.

2. Let $\phi(r,z)$ be the electrostatic potential at a point (r,z) in a situation of axial symmetry, with r the two-dimensional radius such that $r^2 = x^2 + y^2$. Let a be a small quantity and let $r = na$. Show that

$$\phi(r,z) = \{2n[\phi(r,z + a) + \phi(r,z - a)] + (2n + 1)\phi(r + a,z) + (2n - 1)\phi(r - a,z)\}/8n,$$

at a charge-free point in space, correct to order a^3. This is one of the typical theorems useful in the "net point" method of field plotting.

3. Two coaxial pipes of the same diameter with a small gap between them are maintained at a potential difference V. Divide the region within the pipe near the gap into a rectangular net and guess the potentials at each point. Check the correctness of your guesses with the result from the theorem of Exercise 2 above, adjust incorrect values, and repeat until a reasonably correct distribution is obtained. This is called the relaxation method.

4. Find the condition that a set of two-dimensional equipotentials $\phi_2 = f(z,y)$ can generate a set of equipotentials when rotated about the z-axis. Show that if this is possible the potential is

$$\phi = A\int e^{-\int F(\phi_2)d\phi_2}\, d\phi_2 + B$$

where $F(\phi_2) = \frac{1}{y}\frac{1}{(\nabla\phi_2)^2}\frac{\partial\phi_2}{\partial y}$. (See Jeans or Smythe.)

5. Consider the field due to an electric dipole of moment **p**. What charge distribution would have to be introduced on a sphere with **p** at its center to produce zero field outside the sphere?

6. Find a charge distribution that would produce the Yukawa potential

$$\phi = \frac{q}{4\pi\epsilon_0}\frac{e^{-r/a}}{r}.$$

7. Two closed equipotentials ϕ_1, ϕ_0 are such that ϕ_1 contains ϕ_0; ϕ_p is the potential at any point between them. If a charge q is now put at point p and the equipotentials are replaced by grounded conducting surfaces, show that the charges q_1, q_0 induced on the two conductors satisfy the relation

$$q_1/(\phi_0 - \phi_p) = q_0/(\phi_p - \phi_1) = q/(\phi_1 - \phi_0).$$

8. Show that Eqs. (3–31) and (3–37) converge, and that both correspond to the physical situation.

9. Determine the potential inside an infinitely long rectangular prism with grounded conducting walls at $x = 0, a, y = 0, b$, due to a line charge of q per unit length located at the point (c,d) inside the prism.

10. (a) Find the charge density on a grounded spherical conductor of radius a in the presence of a point charge q at a distance $r_0 > a$ from the center of the sphere. (b) Find the charge density on an insulated spherical conductor of radius a in a uniform electric field. (c) What is the least positive charge that must be given

to an insulated spherical conductor of radius a influenced by a point charge q at r_0 in order that the surface charge density be everywhere positive?

11. Obtain the solution of the problem

$$\frac{d^2\phi(x)}{dx^2} = -\frac{\rho(x)}{\epsilon_0}, \quad 0 < x < \pi,$$

$$\phi(0) = A, \quad \phi(\pi) = B,$$

with $\rho(x)$, A, B given arbitrarily, in terms of the Green's function defined by

$$\frac{d^2}{dx'^2} G(x',x) = -\delta(x' - x), \quad 0 < x < \pi$$

$$G(0,x) = G(\pi,x) = 0.$$

CHAPTER 4

TWO-DIMENSIONAL POTENTIAL PROBLEMS

Potential problems involving geometrical arrangements that may be approximated by a two-dimensional geometry with an infinite uniform extent in the third direction are frequently easier to solve than three-dimensional problems. Some methods are really only simpler in two dimensions, and can be generalized to the study of three-dimensional geometry. On the other hand, certain mathematical techniques which have no genuine counterpart for three dimensions may be applied to two-dimensional problems. The method of complex variable potential description combined with conformal transformation is an especially powerful method of this second kind.

4–1 Conjugate complex functions. We shall show that in two dimensions any analytic function W of a complex variable $z_1 = x_1 + iy_1$ will have real and imaginary parts each of which individually satisfies Laplace's equation in two dimensions. (That the symbol z is used for the complex variable should cause no confusion with the third Cartesian coordinate, since all xy-planes are identical in the geometry under consideration.) Thus a suitable function $W = W(z_1)$ can completely describe the potential surfaces and field lines of a particular problem. If $\phi + i\psi = W = W(z_1) = W(x_1 + iy_1)$, we may separate real and imaginary parts and obtain $\phi(x_1,y_1)$ and $\psi(x_1,y_1)$. The equations $\phi =$ constant and $\psi =$ constant will represent the equipotential and field line surfaces or vice versa. Therefore any transformation from one complex variable z_1 to another z_2 will transform the solution of one potential problem described by the first variable to the solution of another potential problem described by the second variable. In general, a whole class of two-dimensional potential distribution problems can be solved by the following process:

(1) Obtain a transformation $z_2 = f(z_1)$ that will transform the geometric arrangement of the z_1 coordinate system into an arrangement of the z_2 coordinate system so as to bring about a simplification in the problem. This coordinate transformation, $z_2 = f(z_1)$ or $z_1 = g(z_2)$, must be so chosen that it will carry the complex potential geometry $W = W_1(z_1)$ of the original problem into a simpler complex potential geometry $W = W_1(g(z_2)) = W_2(z_2)$. The new complex potential W_2 must be that of a more easily soluble problem.

(2) Express the potential solution ϕ in the transformed (i.e., z_2) plane in such a way that $\phi + i\psi$ is an analytic function of a complex variable.

(3) Transform this solution back into the original z_1 plane.

We shall now discuss the justification for this process. Consider a function $W = \phi + i\psi = f(z)$ where $z = x + iy$. If this functional relationship is to be analytic, ϕ and ψ must satisfy the Cauchy-Riemann differential equations:

$$\frac{\partial\psi}{\partial x} = -\frac{\partial\phi}{\partial y}, \tag{4-1}$$

$$\frac{\partial\psi}{\partial y} = \frac{\partial\phi}{\partial x}. \tag{4-2}$$

By partial differentiation of Eq. (4-1) with respect to x and of Eq. (4-2) with respect to y and combination of the two resulting equations,

$$\nabla^2\psi = 0 = \nabla^2\phi. \tag{4-3}$$

The last equality follows from a repetition of the partial differentiation with the roles of x and y interchanged. Thus both ϕ and ψ are harmonic functions.

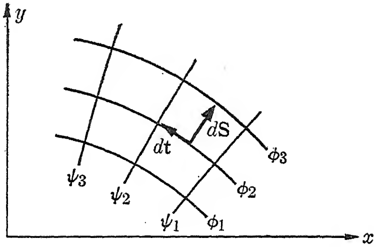


FIG. 4-1 Showing the relation between flux and stream lines.

The functional relationship $W = W(z)$ can be demonstrated graphically (Fig. 4-1) by plotting the lines ϕ = constant and the lines ψ = constant in the $z = x + iy$ plane, after the function W has been separated into its real and imaginary parts. Note that the Cauchy-Riemann relations, Eqs. (4-1) and (4-2), ensure that these curves are normal to each other. The curves of constant ϕ obtained by giving a succession of values to W may be taken to represent the potential field of a problem, and the corresponding ψ curves taken to represent the electric field, although the latter are usually referred to as the streamlines.

The flux of the electric field crossing a surface S may be defined by

$$\Phi = \int \mathbf{E} \cdot d\mathbf{S}. \tag{4-4}$$

Let us consider a surface lying along one of the equipotential curves, ϕ = constant, between two stream curves ψ_1 and ψ_2, and of unit height normal to the xy-plane. For the purposes of this proof let $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ represent unit vectors in the directions of increasing x, y, z, and let $\mathbf{t}$ be the length along the ϕ curve, as shown in Fig. 4-1. Since the surface is of unit height,

$d\mathbf{t} = \mathbf{k} \times d\mathbf{S}$. Then Eq. (4–4) becomes

$$\Phi = \int_1^2 \mathbf{E} \cdot d\mathbf{S} = -\int_1^2 \nabla\phi \cdot d\mathbf{S} = -\int_1^2 \left(\frac{\partial\phi}{\partial x}\mathbf{i} + \frac{\partial\phi}{\partial y}\mathbf{j}\right) \cdot d\mathbf{S}.$$

By substitution from Eqs. (4–1) and (4–2), we have

$$\Phi = \int_1^2 \left(\frac{\partial\psi}{\partial y}\mathbf{i} - \frac{\partial\psi}{\partial x}\mathbf{j}\right) \cdot d\mathbf{S} = \int_1^2 (\nabla\psi \times \mathbf{k}) \cdot d\mathbf{S}$$

$$= \int_1^2 \nabla\psi \cdot d\mathbf{t} = \psi_2 - \psi_1. \tag{4–5}$$

Thus the difference between two stream functions ψ_1 and ψ_2 represents the electric flux passing between the right cylinders of unit height generated by two neighboring lines ψ_1 and ψ_2. This means that no lines of force cross the constant ψ lines. This is the justification for calling ψ the stream function, since in two-dimensional hydrodynamic problems the ψ lines do trace the streamlines of the fluid. In our case, the streamlines traced by giving ψ different constant values will trace the electric field, when ϕ lines are the equipotentials of the field. If, on the other hand, ψ had been assumed to be the potential, then ϕ would have been the stream function. This possibility of exchanging the meaning of ϕ and ψ is frequently useful in the solution of two-dimensional problems.

4–2 Capacity and field strength. The above considerations permit us to obtain immediately the capacity between any two conductors whose boundaries coincide with two equipotential lines ϕ_1 and ϕ_2, and extend between two streamlines ψ_1 and ψ_2. From Eq. (1–26) and the definition of Φ,

$$\Phi = \int \mathbf{E} \cdot d\mathbf{S} = q/\epsilon_0.$$

The capacity is given by

$$C = \frac{q}{|\Delta\phi|} = \frac{\epsilon_0\Phi}{|\Delta\phi|}. \tag{4–6}$$

Since the flux Φ is the change in the stream function ψ between the edges of the conductor surfaces being considered, Eq. (4–5) becomes

$$C = \epsilon_0 \frac{\psi_1 - \psi_2}{\phi_1 - \phi_2}. \tag{4–7}$$

Note that all the charges are assumed to lie on the bounding ϕ surfaces: in general the stream function will be multiple-valued if charges are present in the field.

The absolute magnitude of the field strength can also be calculated from a known function of the form $W = W(z)$ representing a particular geometry as in Fig. 4-1. Consider the modulus of the derivative of W:

$$\left|\frac{dW}{dz}\right| = \left|\frac{\partial(\phi + i\psi)}{\partial x}\frac{dx}{dz} + \frac{\partial(\phi + i\psi)}{\partial y}\frac{dy}{dz}\right|,$$
$$\left|\frac{dW}{dz}\right| = \left|\frac{\partial\phi}{\partial x}dx + i\frac{\partial\psi}{\partial y}dy + i\frac{\partial\psi}{\partial x}dx + \frac{\partial\phi}{\partial y}dy\right|\left|\frac{1}{dz}\right|. \tag{4-8}$$

With the aid of the Cauchy-Riemann equations, we obtain

$$\left|\frac{dW}{dz}\right| = \left|\frac{\partial\phi}{\partial x} - i\frac{\partial\phi}{\partial y}\right| = \sqrt{\left(\frac{\partial\phi}{\partial x}\right)^2 + \left(\frac{\partial\phi}{\partial y}\right)^2} = |\nabla\phi| = |\mathbf{E}|. \tag{4-9}$$

The real and imaginary parts of dW/dz are thus respectively the x- and y-components of the gradient of the potential and therefore the modulus of dW/dz is equal to the magnitude of the electric field strength.

4-3 The potential of a uniform field. Before examining some of the transformations that are useful for simplifying complicated problems, we shall look at two basic potentials from which many more general cases may be generated by transformations and superpositions. The simplest is that of a uniform field $\mathbf{E}$ directed along x, for which $\phi = -|\mathbf{E}|x$. The complex potential can be seen by inspection to be

$$W = -|\mathbf{E}|z = -|\mathbf{E}|(x + iy) = \phi + i\psi \tag{4-10}$$

and the stream function is $\psi = -|\mathbf{E}|y$. (The potential ϕ has been arbitrarily set equal to zero along the y-axis.)

4-4 The potential of a line charge. The Coulomb field around a line charge with a linear charge density q is found by means of Gauss's electric flux theorem, Eq. (1-26), the surface of integration being that of a circular cylinder of radius r and unit height coaxial with the line charge. This field is

$$\mathbf{E} = \frac{q\mathbf{r}}{2\pi\epsilon_0 r^2}. \tag{4-11}$$

The corresponding potential may be secured by substituting this field into Eq. (1-31) and integrating:

$$\phi = -\frac{q}{2\pi\epsilon_0}(\ln r - \ln r_0). \tag{4-12}$$

Again we note that in two-dimensional potential expressions it is not possible to set the potential at infinity equal to zero, since the two-dimensional expression really represents the potentials due to charge distributions of infinite extent perpendicular to the xy-plane. It must be remarked that a two-dimensional problem can be, at most, only an approximation to physical reality, for it implies not only infinite extent but infinite charge. A physical problem can be treated by two-dimensional methods only when it is possible to neglect end effects arising from the finite linear extent of the physical arrangement. Usually it is convenient to choose some conductor as the potential base, whether or not any element of the system is actually grounded. In Eq. (4–12) the cylinder surrounding the line charge at a distance r_0 has been arbitrarily set at zero potential.

The complex potential function W corresponding to the line charge potential may be derived by means of the Cauchy-Riemann equations, but it is easily written down merely by inspection of Eq. (4–12). If we introduce the complex polar notation $z = re^{i\theta}$, we have

$$W = -\frac{q}{2\pi\epsilon_0}(\ln z - \ln z_0) = \frac{-q}{2\pi\epsilon_0}(\ln r + i\theta - \ln r_0) = \phi + i\psi, \tag{4–13}$$

where ϕ represents the potential and ψ represents the stream function. Note that the stream function, as might be expected from the axial symmetry of the arrangement, is proportional to the argument of z, i.e., just the polar angle θ, if z_0 is arbitrarily taken as real. The complex potential function for any system of line charges can be obtained by superposition of appropriately displaced expressions like Eq. (4–13), one for each line charge.

4–5 Complex transformations. We now turn to the analysis of the behavior of curves in a small region of the complex potential plane when a transformation of the plane is made. Consider a transformation from the z_1 plane to the z_2 plane given by the equation $z_2 = f(z_1)$ and let the transformation function f be analytic except at a finite number of singularities. At all nonsingular points such a transformation is conformal. This means that the angle between two intersecting lines in the z_1 plane, such as θ_1 in Fig. 4–2(a), transforms into an equal angle between the transformed lines in the z_2 plane, as θ_2 in Fig. 4–2(b). This can be demonstrated as follows. Since all derivatives of an analytic function of a complex variable exist and are continuous, the derivative dz_2/dz_1 will be finite at all points except for the singularities. Let us consider two line elements intersecting at point P_1, for both of which $dz_2/dz_1 = f'(z_1)$ evaluated at the point P_1. The argument of a product is equal to the sum of the arguments of the factors, so that we have for the argument of the differential line element P_2Q_2:

$$\arg (dz_2) = [\arg f'(z_1)]_{P_1} + \arg (dz_1), \tag{4–14}$$

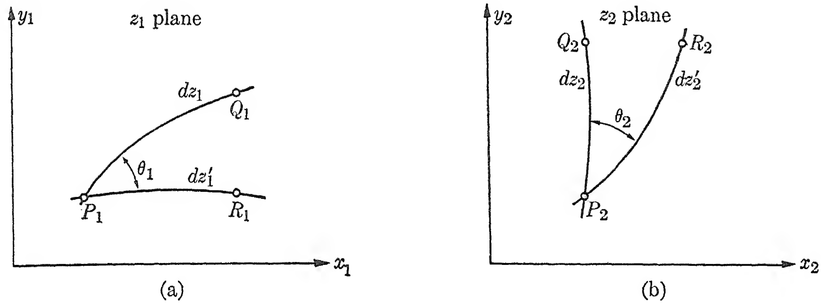


FIG. 4-2 To show that angles are preserved in a complex transformation.

and for the argument of P_2R_2:

$$\arg (dz'_2) = [\arg f'(z_1)]_{P_1} + \arg (dz'_1). \tag{4-15}$$

Subtracting Eq. (4-15) from Eq. (4-14), and noting that the angles θ_1 and θ_2 are the differences in the arguments of the respective dz's, we obtain

$$\theta_1 = \theta_2. \tag{4-16}$$

The modulus of the derivative $dz_2/dz_1 = f'(z_1)$ represents the scale factor by which all spatial intervals in the neighborhood of a point are multiplied. This follows from the fact that the modulus of a product is equal to the product of the moduli of the factors. An infinitesimal triangle will thus transform into a similar infinitesimal triangle in the new system, and

$$|dz_2| = |f'(z_1)_{P_1}| \cdot |dz_1|. \tag{4-17}$$

The similarity of this transformed infinitesimal triangle and the original one provides an alternate way of seeing that angles are preserved in complex transformations. This means that the orthogonality between stream functions and equipotentials is invariant under a complex variable transformation.

4-6 General Schwarz transformation. A transformation that will reduce any number of rectilinear boundaries in the z_1 plane to a single straight line boundary in the z_2 plane is due to Schwarz. The Schwarz transformation will map the inside of a polygon in the z_1 plane (although the polygon need not be closed) into the upper half of the z_2 plane. It is based on the special transformation, useful in itself, which changes the size of an angle whose vertex is at the origin in the z_1 plane, as shown in Fig. 4-3.

Consider the simple transformation

$$z_1 = z_2^{\beta}, \tag{4-18}$$

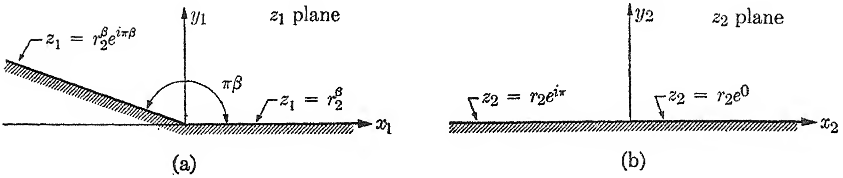

y_1
z_1 plane
$z_1 = r_2^\beta e^{i\pi\beta}$
$\pi\beta$
$z_1 = r_2^\beta$
x_1
(a)
y_2
z_2 plane
$z_2 = r_2 e^{i\pi}$
$z_2 = r_2 e^0$
x_2
(b)

FIG. 4–3 Schwarz transformation for a single angle.

where β is real but not necessarily an integral or a rational number. By this transformation points on the positive real axis are mapped on the positive real axis, although the scale along the axis is changed by raising x_1 to the $1/\beta$ power, or at least a branch of the transformation can be chosen where this is so. On the other hand, for points lying on the negative real axis in the z_2 plane ($z_2 = r_2 e^{i\pi}$), z_1 is complex, since by the transformation z_1 is equal to $r_2^\beta e^{i\pi\beta}$. Hence the negative real axis of the z_2 plane is the mapping of a straight line in the z_1 plane, as required by the conformal properties of the transformation, but this line will make an angle $\pi\beta$ with the positive z_1 axis. The transformation of Eq. (4–18) with $\beta < 1$ therefore maps the area of the upper half of the z_1 plane lying between $\theta_1 = 0$ and $\theta_1 = \pi\beta$ into the entire upper half of the z_2 plane. Of course, β may be greater than one, in which case the angle is obtuse. The transformation has a branch point at $z_2 = 0$, at which angles are obviously not preserved, but it is analytic everywhere else.

In the more general case of Fig. 4–4 there are a number of points b_i in the z_1 plane which are the corners of a polygon whose interior angles are α_i. We wish to map the interior of the polygon into the upper half of the z_2 plane. Consider a transformation defined by the differential equation

$$\frac{dz_1}{dz_2} = C_1 \prod_{i=1}^{n} (z_2 - a_i)^{\beta_i}. \tag{4–19}$$

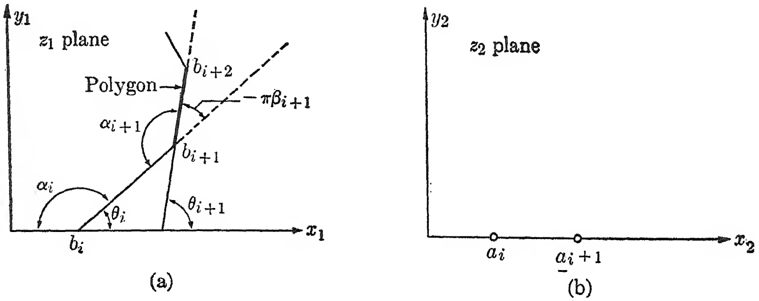

y_1
z_1 plane
Polygon
b_{i+2}
$-\pi\beta_{i+1}$
α_{i+1}
b_{i+1}
α_i
θ_i
θ_{i+1}
b_i
x_1
(a)
y_2
z_2 plane
a_i
a_{i+1}
x_2
(b)

FIG. 4–4 The general Schwarz transformation.

Here C_1 is a constant, possibly complex. This transformation is analytic everywhere except at the points $z_2 = a_i$, which are real but otherwise as yet undetermined. Hence by the conformal properties of such a transformation the real z_2 axis, $z_2 = x_2$, will consist of mappings of straight line segments in the z_1 plane. The angles which each of these straight line segments make with the real axis will be given by the argument of dz_1/dz_2 evaluated in the segment in question. We may take the argument of Eq. (4–19):

$$\arg\left(\frac{dz_1}{dz_2}\right) = \arg C_1 + \beta_1 \arg(z_2 - a_1) + \beta_2 \arg(z_2 - a_2) + \cdots + \beta_n \arg(z_2 - a_n). \tag{4–20}$$

But

$$\arg\left(\frac{dz_1}{dz_2}\right) = \arg\left(\frac{dx_1 + i\,dy_1}{dx_2 + i\,dy_2}\right), \tag{4–21}$$

and when dz_2 lies along the real axis in the z_2 plane $dy_2 = 0$. Therefore Eq. (4–21) becomes, when evaluated in the ith interval,

$$\arg\left(\frac{dz_1}{dz_2}\right) = \arg\left(\frac{dx_1}{dx_2} + i\frac{dy_1}{dx_2}\right) = \tan^{-1}\left(\frac{dy_1}{dx_1}\right) = \theta_i. \tag{4–22}$$

Now when z_2 lies on the real axis between a_i and a_{i+1}, the argument of z_2 minus any point to the left is zero and the argument of z_2 minus any point to the right is just π. Therefore a combination of Eqs. (4–20) and (4–22) leads to

$$\theta_i = \arg C_1 + \pi(\beta_{i+1} + \beta_{i+2} + \cdots \beta_n). \tag{4–23}$$

Thus all points of the real axis segment $a_{i+1} - a_i$ are mappings of a line segment with slope θ_i in the z_1 plane. If we subtract Eq. (4–23) from a similar expression for θ_{i+1}, we obtain

$$\theta_{i+1} - \theta_i = -\pi\beta_{i+1}. \tag{4–24}$$

From the geometry of Fig. 4–4(a) we see that this angle difference of $-\pi\beta_{i+1}$ at the point b_{i+1} is related to the interior angle α_{i+1} at each corner by the relation

$$\alpha_{i+1} = \pi + \pi\beta_{i+1}. \tag{4–25}$$

For convenience, we may change the subscript to i and solve for β_i,

$$\beta_i = \frac{\alpha_i}{\pi} - 1. \tag{4–26}$$

Hence Eq. (4–19) becomes

$$\frac{dz_1}{dz_2} = C_1 \prod_{i=1}^{n} (z_2 - a_i)^{\left(\frac{\alpha_i}{\pi} - 1\right)}, \tag{4–27}$$

where the scale factor C_1 gives both the relative scale and the relative angular orientation of the two geometries. This is the required transformation for a polygon with internal angles α_i.

In general, the Schwarz transformation is a useful one provided that Eq. (4–27) is integrable in terms of elementary functions. This is possible, with the exception of special cases, only when the angles are multiples of 90° and not more than two corners are involved. One further difficulty in the practical application of the Schwarz transformation is that the resultant transformation is given in terms of $z_1 = f(z_2)$ with the coordinate along the real axis in the z_2 plane as the independent variable, rather than in terms of the coordinates of the z_1 plane, or the problem polygon, as the independent variables. Therefore considerable computational labor is often necessary to find out what the coordinates a_i in the z_2 plane actually are in terms of the geometry of the given problem. Once the a_i are determined, the remainder of the solution of the potential problem is usually simple. The consideration of some special simple cases will serve to indicate the power of the method and to illustrate its use.

4–7 Single-angle transformations. A single angle can always be transformed into the origin of the z_2 plane, so that $a_i = 0$ in each case. For some special angles Eq. (4–27) is immediately integrable:

(a) $\alpha = \pi$. The integration of Eq. (4–27) gives

$$z_1 = C_1 z_2 + C_2. \tag{4–28}$$

This is simply a uniform translation and rotation, and is of no physical interest.

(b) $\alpha = \pi/2$. The integration of Eq. (4–27) gives

$$z_1 = C_3 z_2^{1/2} + C_2. \tag{4–29}$$

If we assume that the constant of translation C_2 is zero, Eq. (4–29) will map the first quadrant of the z_1 plane into the upper half of the z_2 plane. If, for example, the complex potential in the z_2 plane is given by the complex potential solution corresponding to a uniform field $\mathbf{E}$,

$$\begin{aligned} W &= -|\mathbf{E}|z_2 = \phi + i\psi, \\ \phi &= -|\mathbf{E}|x_2, \quad \psi = -|\mathbf{E}|y_2, \end{aligned} \tag{4–10}$$

then W in the z_1 plane, according to Eq. (4–29), is

$$\begin{aligned} W &= -C_4|\mathbf{E}|z_1^2 \\ &= -C_4|\mathbf{E}|(x_1^2 - y_1^2 + 2ix_1y_1), \\ \phi &= -C_4|\mathbf{E}|(x_1^2 - y_1^2), \\ \psi &= -C_4|\mathbf{E}|(2x_1y_1). \end{aligned}$$

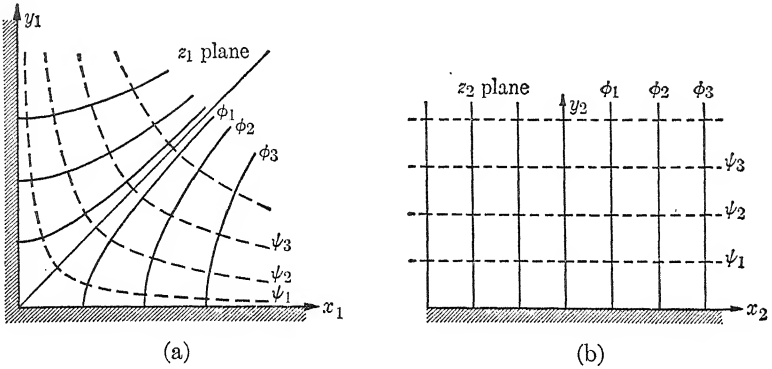


FIG. 4-5 Schwarz transformation for $\alpha = \pi/2$.

This will solve the problem of the charged rectangular boundary (see Fig. 4-5) or, if the interpretations of ϕ and ψ are interchanged, problems involving charged hyperbolic cylinders.

If the complex potential in the z_2 plane is taken to be the logarithmic potential corresponding to a line charge, Eq. (4-12), and if the transformation of Eq. (4-29) is then applied, we obtain the two-dimensional Green's function for an inside rectangular corner, provided that we have translated the line charge into the upper half of the z_2 plane. This same transformation will give the Green's function for a problem having hyperbolic cylindrical boundaries, and thus problems involving such geometries are amenable to solution.

(c) $\alpha = 0$. In this case, the integration yields

$$z_1 = C_3 \ln z_2 = C_3 \ln r_2 + C_3 i\theta_2, \tag{4-30}$$

if we omit the translation constant. If C_3 is real the real part of z_1 is $C_3 \ln r_2$, the positive real z_2 axis is the mapping of the whole real z_1 axis, and the upper half z_2 plane maps into a strip of width $C_3\pi$, as in Fig. 4-6. The transformation can be visualized by considering the origin in the z_2

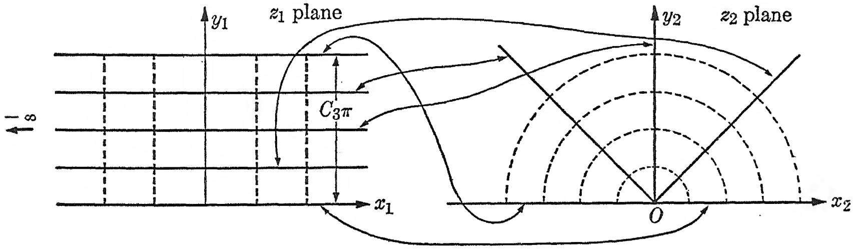


FIG. 4-6 Schwarz transformation for $\alpha = 0$.

plane to be pushed to minus infinity, and the negative real axis of the z_2 plane to be revolved clockwise to a position parallel to the positive real axis but located above it a distance $C_3\pi$, as seen in Fig. 4–6. This transformation thus results in a periodic configuration in the z_1 plane; of this the strip $C_3\pi$ wide is the first repeat. The upper half of the z_2 plane is the mapping of the first strip of this configuration in the z_1 plane. The lower half of the z_2 plane is the mapping of the strip lying between $y_1 = C_3\pi$ and $y_1 = 2C_3\pi$ in the z_1 plane, and so on. This transformation is a very useful one in the solution of potential problems involving grids, repeating condenser plates, and other geometries that repeat in one direction.

4–8 Multiple-angle transformations. A simple example is one of the most useful. If two vertical lines in the z_1 plane are rotated into the positive and negative real axes, respectively, then the upper half of the z_2 plane will be the mapping of the vertically oriented, semi-infinite strip seen in Fig. 4–7(a). The two transformed corners may be taken at $z_2 = \pm a$, and the differential equation becomes

$$\frac{dz_1}{dz_2} = C_1(z_2 - a)^{-1/2}(z_2 + a)^{-1/2} = \frac{C_1}{\sqrt{z_2^2 - a^2}} = \frac{C_2}{\sqrt{a^2 - z_2^2}}. \tag{4–31}$$

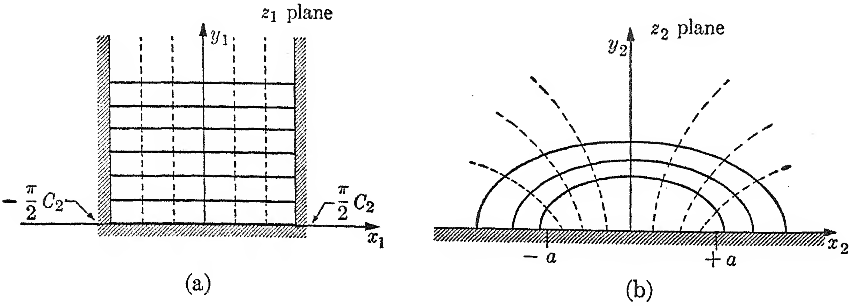
The relation $C_1 = iC_2$ has been introduced to rotate the figure 90° to the orientation shown. Integration yields

$$z_1 = C_2 \sin^{-1}\left(\frac{z_2}{a}\right), \quad z_2 = a \sin\left(\frac{z_1}{C_2}\right). \tag{4–32}$$

In practice this transformation is most often used to transform a solution in the z_1 plane into the z_2 plane. If we consider a uniform complex potential field W in the z_1 plane,

$$W = -|\mathbf{E}|z_1, \tag{4–10}$$

mapped into the z_2 plane by Eq. (4–32), we see that we have obtained the cross section of the potential around a charged conducting strip of



(a)

(b)

FIG. 4–7 Two-angle transformation.

width $2a$, or, exchanging potential and stream functions, the potential due to a slot in a conducting sheet. The major axis of the slot or strip is normal to the plane of Fig. 4-7. If the real and imaginary parts of the transformed W are given sets of constant values they will characterize the field of the arrangement. These equations turn out to be the equations of confocal elliptic and hyperbolic cylinders, as indicated in Fig. 4-7.

Many other practical examples of double-angle transformations are readily integrable, and still others may be most easily found by simple successive transformations, as indicated in the references listed at the end of this chapter. Frequently the solution appears in a form that is deceptively simple, since, as we have noted earlier, it is sometimes very difficult to solve for the z_2 coordinates as a function of z_1. Despite these difficulties, the method is obviously a powerful one.

4-9 Direct solution of Laplace's equation by the method of harmonics. Many two-dimensional problems are at least as conveniently solvable by methods for which analogs do exist in three dimensions as by complex potential methods. Two-dimensional inversions and images are useful special cases of general methods already treated. The solution of the two-dimensional Laplace equation by separation of variables for plane polar coordinates is particularly useful, since it has general validity whenever circular or radial boundaries are encountered. Let us consider the application of this method in some detail.

Laplace's equation in the plane polar coordinates r and θ is

$$r\frac{\partial}{\partial r}\left(r\frac{\partial\phi}{\partial r}\right) + \frac{\partial^2\phi}{\partial\theta^2} = 0. \tag{4-33}$$

To achieve separation we let $\phi(r,\theta) = R(r)\Theta(\theta)$, substitute in Eq. (4-33) and divide by ϕ. The result is

$$\frac{r}{R}\frac{\partial}{\partial r}\left(r\frac{\partial R}{\partial r}\right) + \frac{1}{\Theta}\frac{\partial^2\Theta}{\partial\theta^2} = 0. \tag{4-34}$$

The two terms must be individually constant, and we may choose the sign of the separation parameter so as to give circular functions in the angle θ. Since the maximum range of θ is 2π, the boundaries of θ are always "closed," so to speak, and only certain values of the parameter will be allowed, just as in the first treatment of the example of Section 3-6. In other words, we may set the first term of Eq. (4-34) equal to k_n^2, to obtain

$$\begin{gathered}\frac{\partial^2\Theta}{\partial\theta^2} + k_n^2\Theta = 0,\\ r\frac{\partial}{\partial r}\left(r\frac{\partial R}{\partial r}\right) - k_n^2R = 0\end{gathered} \tag{4-35}$$

where the separation parameter k_n^2 is in general restricted to discrete values. For $k_n \neq 0$ the solutions are

$$\Theta_n = A_n \cos k_n\theta + B_n \sin k_n\theta,$$

$$R_n = C_n r^{k_n} + D_n r^{-k_n},$$

and if $k_n = 0$,

$$\Theta_0 = E + F\theta,$$

$$R_0 = G + H \ln r.$$

The general solution is obtained by a linear superposition of the individual ("harmonic") solutions:

$$\begin{aligned}\phi &= \sum_{n=0}^{\infty} R_n\Theta_n \\ &= \sum_{n=1}^{\infty} (A_n \cos k_n\theta + B_n \sin k_n\theta)(C_n r^{k_n} + D_n r^{-k_n}) \\ &\qquad + (E + F\theta)(G + H \ln r).\end{aligned} \tag{4-36}$$

In order to apply Eq. (4-36) to the solution of a practical problem, we shall first express certain already known potentials in this form, and then superpose additional potentials, with undetermined coefficients, of the same general form. The coefficients are determined by the boundary conditions of the given problem, just as in Section 3-6.

4-10 Illustration: Line charge and dielectric cylinder. To solve the problem of a line charge located at a distance r_0 from the axis of a dielectric cylinder of radius a and specific inductive capacity κ, as seen in Fig. 4-9, let us first express the logarithmic potential of the line charge alone in the same form as Eq. (4-36). This amounts to shifting the origin of the potential, as indicated in Fig. 4-8. We may omit the arbitrary potential base, so that Eq. (4-12) reduces to

$$\phi = -\frac{q}{2\pi\epsilon_0} \ln R. \tag{4-37}$$

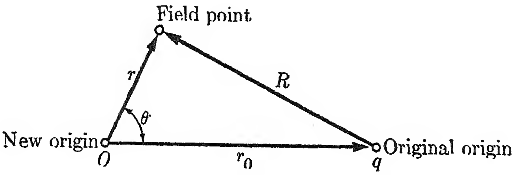


FIG. 4-8 Shifting the origin of coordinates for the potential.

Since Eq. (4–36) is in general nonsingular except at the origin $r = 0$, Eq. (4–37) cannot in general be expressed by a single expansion about the new origin, but must be written as two different solutions, one valid in the region where $r < r_0$ and one valid where $r > r_0$. These two solutions must fit together at $r = r_0$ in such a way that the derivative shall be discontinuous only at the point where the line charge is located, and continuous at all other points. The discontinuity is such that the total flux emerging from that point corresponds to the value of the line charge per unit length.

The logarithmic potential Eq. (4–37) of a line charge at the origin can be put into the form of Eq. (4–36) of an isolated line charge located at $\theta = 0$ and $r = r_0$, as in Fig. 4–8, if we express the radial distance R by the law of cosines, $R = (r^2 + r_0^2 - 2rr_0 \cos\theta)^{1/2}$, and then expand in a power series in r/r_0 for use where $r < r_0$, and in a power series in r_0/r to use where $r > r_0$. Both series converge within their respective ranges of validity. The result gives us the potential due to a line charge only:

$$\underset{0<r<r_0}{\phi} = \frac{q}{2\pi\epsilon_0}\left\{\sum_{n=1}^{\infty} \frac{1}{n}\left(\frac{r}{r_0}\right)^n \cos n\theta - \ln r_0\right\}, \tag{4–38}$$

$$\underset{r_0<r<\infty}{\phi} = \frac{q}{2\pi\epsilon_0}\left\{\sum_{n=1}^{\infty} \frac{1}{n}\left(\frac{r_0}{r}\right)^n \cos n\theta - \ln r\right\}. \tag{4–39}$$

For the problem of the line charge and dielectric cylinder we shall choose the origin of the polar coordinate system at the center of the dielectric cylinder, with the radius vector corresponding to $\theta = 0$ passing through the charge, as shown in Fig. 4–9. This is the same coordinate system as that in which Eqs. (4–38) and (4–39) are written. To satisfy the boundary condition at the surface of the cylinder $r = a$ we shall superpose on

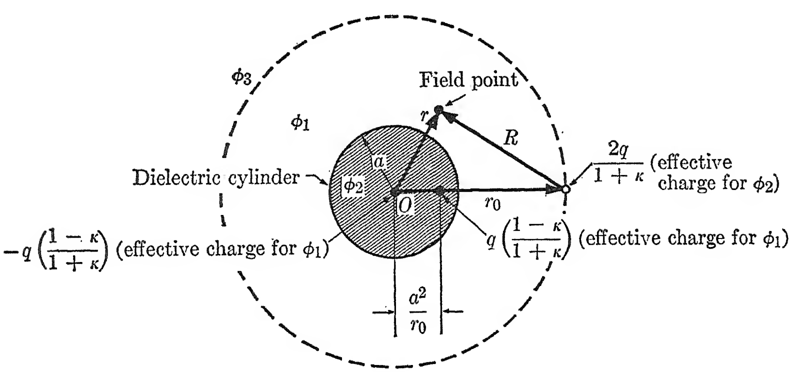

ϕ_3
Field point
ϕ_1
r
R
a
Dielectric cylinder
ϕ_2
O
r_0
$\frac{2q}{1+\kappa}$ (effective charge for ϕ_2)
$q\left(\frac{1-\kappa}{1+\kappa}\right)$ (effective charge for ϕ_1)
$-q\left(\frac{1-\kappa}{1+\kappa}\right)$ (effective charge for ϕ_1)
$\frac{a^2}{r_0}$

FIG. 4–9 Line charge and dielectric cylinder.

the line charge solution (4–38) a general solution of the type of Eq. (4–36) with $k_n = n$ and with undetermined coefficients A_n, B_n, E, and F, and make a separation of the potential into two parts ϕ_1 and ϕ_2 to be valid outside and inside the cylinder respectively. The coefficients are to be determined so as to account for the polarization of the dielectric cylinder, and the separation is made to ensure a finite value for the potential at the origin and convergence of the second series in each expression.

$$\underset{a<r<r_0}{\phi_1} = \frac{q}{2\pi\epsilon_0}\left\{\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{r}{r_0}\right)^n \cos n\theta - \ln r_0\right\} + \sum_{n=1}^{\infty} B_n r^{-n}\cos n\theta + F, \tag{4–40}$$

$$\underset{0<r<a}{\phi_2} = \frac{q}{2\pi\epsilon_0}\left\{\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{r}{r_0}\right)^n \cos n\theta - \ln r_0\right\} + \sum_{n=1}^{\infty} A_n r^{n}\cos n\theta + E. \tag{4–41}$$

Since the effect of the induced polarization charges in the cylinder is nonsingular both at the origin and at infinity, and the solution is obviously symmetric about the line $\theta = 0$, it has already been possible to simplify the expression (4–36) by omitting the logarithm, the angle and its sine, and by using only positive or negative powers of r as necessary to assure convergence.

For any angle θ the boundary conditions of Section 2–2 for the surface of the dielectric $r = a$ are just

$$\phi_1 = \phi_2, \quad \frac{\partial\phi_1}{\partial r} = \kappa\frac{\partial\phi_2}{\partial r}. \tag{4–42}$$

We can evaluate the coefficients A_n, B_n, E, and F by applying these conditions to Eqs. (4–40) and (4–41) and then equating the coefficients of $\cos n\theta$, term by term, to zero. This procedure is justified mathematically by the fact that these Fourier series functions form a complete orthogonal set. The resultant solution is

$$\underset{a<r<r_0}{\phi_1} = \frac{q}{2\pi\epsilon_0}\left\{\sum_{n=1}^{\infty}\frac{1}{n}\left[\left(\frac{r}{r_0}\right)^n + \left(\frac{1-\kappa}{1+\kappa}\right)\left(\frac{a^2}{r_0}\right)^n\frac{1}{r^n}\right]\cos n\theta - \ln r_0\right\}, \tag{4–43}$$

$$\underset{0<r<a}{\phi_2} = \frac{q}{\pi\epsilon_0(1+\kappa)}\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{r}{r_0}\right)^n\cos n\theta - \frac{q}{2\pi\epsilon_0}\ln r_0. \tag{4–44}$$

For r greater than r_0 the solution may be written down immediately, analogously to Eqs. (4–38) and (4–39).

If we add zero to Eq. (4–43) by adding and subtracting

$$q\,\frac{(1-\kappa)}{(1+\kappa)}\,\frac{\ln r}{2\pi\epsilon_0},$$

we see that the potential ϕ_1 outside the cylinder corresponds to an effective line charge arrangement, with the role of the cylinder taken by two line charges. This arrangement consists of an effective charge $-q(1-\kappa)/(1+\kappa)$ located at the origin, an effective charge $q(1-\kappa)/(1+\kappa)$ located at the inversion point of the actual external charge, and the actual charge. The inversion point lies on the vector r_0 at a distance a^2/r_0 from the origin. This problem could therefore have been solved by the method of images, a fact which can be verified directly by the use of the logarithmic potentials. On the other hand, the potential ϕ_2 inside the cylinder is seen to correspond to a single effective charge at the position of the actual charge but of strength $2q/(1+\kappa)$, except for an additive constant. The potential everywhere, then, is equivalent to some line charge arrangement and the dielectric cylinder absent.

4–11 Line charge in an angle between two conductors.

As another example of the solution of a problem in terms of circular harmonics, let us consider a wedge-shaped region bounded by grounded conducting surfaces intersecting at the origin with an interior angle α, as in Fig. 4–10, together with a line charge of strength q per unit length located at the point r_0, β within the wedge. The solution of this problem will give the Green's function for the region bounded by the intersecting conducting planes. It is again clear that we cannot hope to express a solution by means of a single equation valid throughout the region from $r = 0$ to $r = \infty$. The solution must be written in two parts, one valid for $r < r_0$ and the other valid for $r > r_0$, joined together at the cylindrical surface $r = r_0$ by the flux condition corresponding to the charge q.

Since the potential vanishes on the boundaries $\theta = 0$ and $\theta = \alpha$, the angular part of the solution must be of the form $\sin(n\pi\theta/\alpha)$. Thus, in

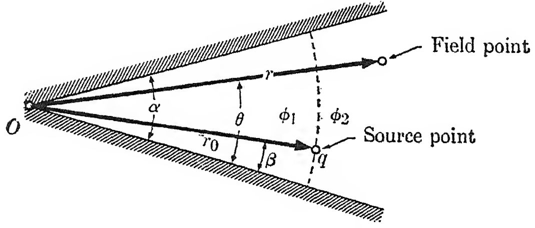


FIG. 4–10 Line charge parallel to two intersecting plane conductors.

Eq. (4–36), $k_n = n\pi/\alpha$. The potentials are then formally:

$$\begin{aligned} \underset{r<r_0}{\phi_1} &= \sum_{n=1}^{\infty} C_n \left(\frac{r}{r_0}\right)^{n\pi/\alpha} \sin\frac{n\pi}{\alpha}\theta, \\ \underset{r>r_0}{\phi_2} &= \sum_{n=1}^{\infty} D_n \left(\frac{r_0}{r}\right)^{n\pi/\alpha} \sin\frac{n\pi}{\alpha}\theta. \end{aligned} \tag{4–45}$$

The coefficients C_n and D_n must be equal in order to assure continuity of the potentials across the cylindrical surface $r = r_0$, so that we shall have solved the problem if we determine the C_n.

The line charge is equivalent to a surface charge $\sigma(\theta)$ on the cylinder $r = r_0$, where $\sigma(\theta) = q\delta(\theta - \beta)$. The angular δ-function can be defined by the equations

$$\begin{aligned} r_0\int_0^{\alpha} \delta(\theta - \beta)\, d\theta &= 1, \\ \int_0^{\alpha} f(\theta)\, \delta(\theta - \beta)\, d\theta &= f(\beta)/r_0, \quad 0 < \beta < \alpha. \end{aligned} \tag{4–46}$$

The flux condition is thus

$$q\,\delta(\theta - \beta) = -\epsilon_0\left[\frac{\partial\phi_2}{\partial r} - \frac{\partial\phi_1}{\partial r}\right]_{r=r_0}. \tag{4–47}$$

The determination of the coefficients is almost identical with that of Section 3–6, and leads to

$$C_n = \frac{q}{n\pi\epsilon_0}\sin\left(\frac{n\pi\beta}{\alpha}\right). \tag{4–48}$$

The complete solution is therefore

$$\begin{aligned} \underset{r<r_0}{\phi_1} &= \frac{q}{\pi\epsilon_0}\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{r}{r_0}\right)^{n\pi/\alpha}\sin\frac{n\pi\beta}{\alpha}\sin\frac{n\pi\theta}{\alpha}, \\ \underset{r>r_0}{\phi_2} &= \frac{q}{\pi\epsilon_0}\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{r_0}{r}\right)^{n\pi/\alpha}\sin\frac{n\pi\beta}{\alpha}\sin\frac{n\pi\theta}{\alpha}. \end{aligned} \tag{4–49}$$

Just as in the case of parallel plates, this solution is the desired Green's function when q is set equal to unity. It is thus evident that the method is a general one for deriving the Green's function within a set of boundaries corresponding to equi-coordinate planes.

SUGGESTED REFERENCES

W. R. SMYTHE, *Static and Dynamic Electricity.* Chapter IV is devoted to two-dimensional potential problems, with many examples.

J. C. MAXWELL, *Electricity and Magnetism,* Vol. 1. There have been many elaborations of potential theory since Maxwell's time, but this remains one of the best fundamental texts.

J. H. JEANS, *The Mathematical Theory of Electricity and Magnetism.* Follows the general presentation of Maxwell, and gives many examples.

E. WEBER, *Electro-Magnetic Fields,* Vol. 1. A more modern treatment.

L. A. PIPES, *Applied Mathematics for Engineers and Physicists.* Chapter XX is a good example of the clear and simple treatment of conjugate functions found in several books on applied mathematics.

EXERCISES

1. By inversion, find the law of image formation for a line charge parallel to a conducting circular cylinder. Apply this method to the case of a large cylinder of radius b containing a smaller cylinder of radius a, the distance between their axes being $c < b - a$. Find the capacity per unit length, and show that for $c = 0$ the result reduces to that for coaxial cylinders.

2. Derive the two-dimensional form of Green's boundary value theorem: if $\phi(x,y)$ is the two-dimensional potential, show that

$$\phi(x',y') = \frac{1}{2\pi}\left\{\int_S (\log r)\nabla^2\phi\, dS - \oint_c \left(\log r\frac{\partial\phi}{\partial n} - \phi\frac{\partial(\log r)}{\partial n}\right) dl\right\},$$

where S is the area bounded by the contour C and r is the distance between the point x, y and the point x', y'. n is the outward normal.

3. Find the field surrounding a charged conducting cylinder whose cross section is an ellipse.

4. Consider two planes intersecting at right angles raised to potentials $V/2$ and $-V/2$, respectively. Calculate the electrostatic field.

5. The transformation

$$z_2 = \frac{a}{2}\left(\frac{z_1}{a_1} + \frac{a_1}{z_1}\right)$$

transforms the region outside the cylinder $r_1 = a_1$ in the z_1 plane into the entire z_2 plane with a cut along the real axis for $-a < x < a$. By transforming the complex potential function corresponding to a conducting cylinder in a uniform field, find the complex potential function corresponding to a strip of width $2a$ with its plane (a) in the direction of an applied field **E**, and (b) perpendicular to an applied field **E**.

6. Consider a parallel plate condenser, $y = \pm a$, of infinite extent, and $y = 0$ semi-infinite, as indicated in Fig. 4-11. The two outside plates are at the same potential. Calculate the capacity per unit length along z contributed by the edge effect, i.e., the difference between the actual capacity per unit length and that estimated by assuming zero field for $x < 0$ and uniform fields for $x > 0$.

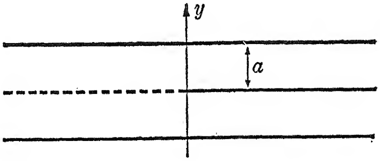


FIG. 4-11

7. An infinite circular cylindrical sheet of radius a is divided longitudinally into quarters which are raised to potentials V, 0, $-V$, 0, respectively. Show that the potential inside the cylinder is given by

$$\phi = \frac{V}{\pi}\left\{\tan^{-1}\left(\frac{2ay}{a^2 - r^2}\right) + \tan^{-1}\left(\frac{2ax}{a^2 - r^2}\right)\right\}.$$

What is the potential outside?

8. Consider the region of space between the cylinder $x^2 + y^2 = b^2$ and the xz-plane. All the curved boundary and that portion of the plane boundary for which $a < x < b$ are at zero potential. That portion of the plane boundary for which $-a < x < a$ is at potential V. Show that the equation of the lines of force in the region for which $a < r < b$ is

$$\sum_n \frac{a^n}{n}\left[\frac{1}{r^n} + \left(\frac{r}{b^2}\right)^n\right] \cos n\theta = \text{constant},$$

where only odd values of n are taken.

9. Let $\phi(x,y)$ be the potential in a two-dimensional field. Let $F(x) = \phi(x,0)$ and F_n be the nth partial derivative of F with respect to x. Show that, if $\phi(x,y) = \phi(x,-y)$,

$$\phi(x,y) = \Sigma A_n y^{2n} F_{2n},$$

where $A_n = (-1)^n/(2n)!$.

CHAPTER 5

THREE-DIMENSIONAL POTENTIAL PROBLEMS

Laplace's equation may readily be written in any orthogonal coordinate system for which the infinitesimal line elements are known: the operator ∇^2 is just the divergence of the gradient, and the application of the divergence theorem to the gradient over the surface of an infinitesimal volume element yields the required expression.* In a number of these systems the equation is separable, so that the methods of Section 3–6 can be applied. The two coordinate systems treated in this chapter further illustrate the general method, and correspond to geometrical configurations often encountered in practice, namely, spheres and circular cylinders, or parts thereof. Spherical coordinates also furnish a particularly useful representation of the potential due to an arbitrary distribution of charge confined to a region near the origin of coordinates.

5–1 The solution of Laplace's equation in spherical coordinates.

Expressed in spherical polar coordinates, Laplace's equation becomes

$$\nabla^2\phi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\phi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\phi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\phi}{\partial\varphi^2} = 0. \tag{5–1}$$

In order to separate the radial and angular parts of this equation, we put

$$\phi = R(r)Y(\theta,\varphi) \tag{5–2}$$

and proceed in the usual way. The separated equations are

$$\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right) - n(n+1)R = 0, \tag{5–3}$$

$$\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{1}{\sin\theta}\frac{\partial^2 Y}{\partial\varphi^2} + n(n+1)\sin\theta Y = 0. \tag{5–4}$$

The form of the separation constant $n(n+1)$, where n is a real integer, is dictated by the necessity that there be a regular solution at the singularities of the equation for Y, which occur at $\theta = 0$ and $\theta = \pi$. In general, $Y(\theta,\varphi)$ is known as a spherical harmonic. It has already been proved that the set of functions $Y_n(\theta,\varphi)$ have orthogonality properties similar to those of the Fourier functions we have considered in connection with two-dimensional solutions. The general solution of the differential equation for the radial part of ϕ, Eq. (5–3), is simply

---

* See Appendix II.

$$R(r) = A_n r^n + B_n r^{-n-1}. \tag{5–5}$$

The spherical surface harmonics can be further separated by means of the substitution

$$Y(\theta,\varphi) = \Theta(\theta)\Phi(\varphi). \tag{5–6}$$

If the new separation constant is called m^2, then, in terms of the more convenient polar angle variable, $\mu = \cos\theta$, the two resulting equations are

$$\frac{d}{d\mu}\left[(1-\mu^2)\frac{d\Theta}{d\mu}\right] + \left[n(n+1) - \frac{m^2}{1-\mu^2}\right]\Theta = 0, \tag{5–7}$$

$$\frac{d^2\Phi}{d\varphi^2} + m^2\Phi = 0. \tag{5–8}$$

The solutions of these equations are

$$\Theta = C_n P_n^m(\mu) + D_n Q_n^m(\mu). \tag{5–9}$$

$$\left.\begin{aligned} \Phi &= E_m \cos m\varphi + F_m \sin m\varphi, & m &\neq 0, \\ \Phi &= G\varphi + H, & m &= 0. \end{aligned}\right\} \tag{5–10}$$

The functions $P_n^m(\cos\theta)$ and $Q_n^m(\cos\theta)$ are the associated Legendre functions of the first and second kind, respectively. Their mathematical properties can be found in numerous references, some of which are listed at the end of the chapter. We need note here only that P_n^m is the solution that is finite for $\mu = \pm 1$, and thus the only solution allowed when the space involved in the problem includes the polar axis.

5–2 The potential of a point charge. For problems having azimuthal symmetry, so that the solution does not depend on the value of the coordinate φ, the separation parameter m is equal to zero. The potential of a point charge at a distance r_0 from the origin of coordinates has such symmetry if the radius vector of the charge is taken as the polar axis. We may obtain the potential of a point charge, expressed in terms of a series expansion in the radial and angular functions obtained in the above separation of coordinates, by expanding the cosine law expression for $1/R$ (see Fig. 5–1) in powers of r/r_0 and r_0/r.

$$\frac{1}{R} = \frac{1}{r_0}\left[\left(\frac{r}{r_0}\right)^2 + 1 - 2\frac{r}{r_0}\cos\theta\right]^{-1/2} = \frac{1}{r}\left[\left(\frac{r_0}{r}\right)^2 + 1 - 2\frac{r_0}{r}\cos\theta\right]^{-1/2} \tag{5–11}$$

becomes, according to the region of convergence,

$$\frac{1}{R} = \frac{1}{r_0}\sum_{n=0}^{\infty}\left(\frac{r}{r_0}\right)^n P_n(\mu)$$

$$= \frac{1}{r}\sum_{n=0}^{\infty}\left(\frac{r_0}{r}\right)^n P_n(\mu). \qquad (5\text{–}12)$$

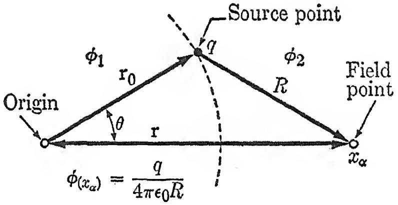

Source point
q
ϕ_1
ϕ_2
r_0
R
Field point
Origin
θ
r
x_α
$\phi(x_\alpha) = \frac{q}{4\pi\epsilon_0 R}$

FIG. 5–1 New polar coordinates for the potential due to a point charge.

Equation (5–12) is often taken as the definition of $P_n(\mu)$. It is obvious from carrying out the expansion that these functions are polynomials of degree n in the variable μ. By explicit differentiation of $1/R$ with respect to μ and with respect to the expansion variable, it can be readily shown that $P_n(\mu)$ satisfies Eq. (5–7) with $m = 0$.

For the same reasons as in the harmonic expansions of Chapter 4, we must use two potentials, one valid where r is less than r_0 and one for r greater than r_0. The two potentials are

$$\underset{r<r_0}{\phi_1} = \frac{q}{4\pi\epsilon_0 r_0}\sum_{n=0}^{\infty}\left(\frac{r}{r_0}\right)^n P_n(\mu),$$

$$\underset{r>r_0}{\phi_2} = \frac{q}{4\pi\epsilon_0 r}\sum_{n=0}^{\infty}\left(\frac{r_0}{r}\right)^n P_n(\mu). \qquad (5\text{–}13)$$

The resulting potential of the point charge is therefore just a Taylor-Laurent series in r whose coefficients are the Legendre polynomials in $\cos\theta$.

5–3 The potential of a dielectric sphere and a point charge. Problems involving point charges and boundaries which have spherical symmetry can be solved in terms of the functions of the foregoing section. As an example, we shall consider the simple problem of a point charge and a dielectric sphere of radius a, as shown in Fig. 5–2, with r_0 being the distance from the center of the sphere to the point charge. We shall need

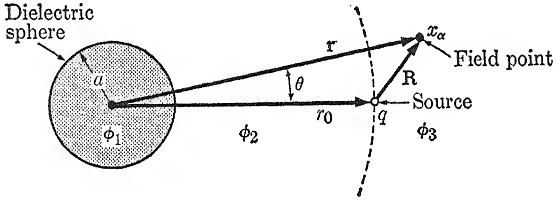

Dielectric sphere
a
ϕ_1
r
θ
r_0
ϕ_2
x_α
Field point
R
Source
q
ϕ_3

FIG. 5–2 Dielectric sphere and point charge.

three expressions for ϕ:

$$
\begin{aligned}
\underset{0<r<a}{\phi_1} &= \sum_{n=0}^{\infty} A_n r^n P_n(\mu),\\
\underset{a<r<r_0}{\phi_2} &= \frac{q}{4\pi\epsilon_0 r_0}\sum_{n=0}^{\infty}\left(\frac{r}{r_0}\right)^n P_n(\mu) + \sum_{n=0}^{\infty} B_n r^{-n-1} P_n(\mu), \qquad (5\text{–}14)\\
\underset{r_0<r<\infty}{\phi_3} &= \frac{q}{4\pi\epsilon_0 r}\sum_{n=0}^{\infty}\left(\frac{r_0}{r}\right)^n P_n(\mu) + \sum_{n=0}^{\infty} B_n r^{-n-1} P_n(\mu).
\end{aligned}
$$

The fit of ϕ_1 to ϕ_2 at $r = a$ can be made in the same way as for the two-dimensional case. The boundary conditions of Eqs. (2–15) and (2–19) must be satisfied for all values of the angle θ, and the fact that the angular functions are orthogonal makes it possible to equate the terms of the series separately to determine the coefficients A_n and B_n. The fit of ϕ_2 to ϕ_3 is inherent from the nature of the solutions in Eq. (5–13). For $\kappa = \infty$ the solution reduces to that obtained by the inversion process outlined in Chapter 3.

5–4 The potential of a dielectric sphere in a uniform field. As a second example of a problem with spherical geometry, let us consider a dielectric sphere of specific inductive capacity κ in the presence of a uniform field whose force lines are parallel to the x-axis, as shown in Fig. 5–3. The lines of electric displacement **D** are shown.

Since the field at infinity is uniform, the potential is given by

$$\phi_\infty = -E_0 x = -E_0 r\cos\theta = -E_0 r\mu = -E_0 r P_1(\mu). \qquad (5\text{–}15)$$

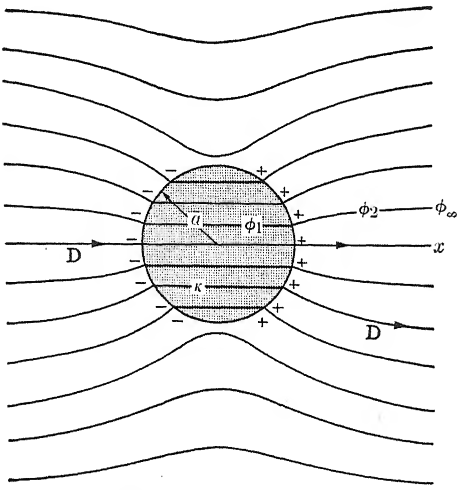


FIG. 5–3 Dielectric sphere in a uniform field.

By inspection of Eqs. (5–5) and (5–9), we may write

$$\begin{aligned}\phi_1 &= \Sigma A_n r^n P_n(\mu),\\ \phi_2 &= \sum_{n=0}^{\infty} B_n r^{-n-1} P_n(\mu) - E_0 r P_1(\mu),\end{aligned} \tag{5–16}$$

for the potentials inside and outside the sphere. The boundary conditions, $\phi_1 = \phi_2$ and $\kappa(\partial\phi_1/\partial r) = (\partial\phi_2/\partial r)$ at $r = a$, must hold for all values of the angle θ. We therefore evaluate the constants A_n and B_n by equating the coefficients of $P_n(\mu)$ for the same n, and find that

$$\begin{aligned} A_0 &= B_0 = 0 = A_n = B_n \quad \text{for } n \text{ greater than } 1,\\ A_1 &= \frac{-3E_0}{\kappa+2},\\ B_1 &= \frac{(\kappa-1)E_0 a^3}{\kappa+2}.\end{aligned} \tag{5–17}$$

The potentials are therefore

$$\begin{aligned}\phi_1 &= -\frac{3E_0 r}{\kappa+2}\cos\theta,\\ \phi_2 &= \frac{(\kappa-1)}{(\kappa+2)}\frac{E_0 a^3\cos\theta}{r^2} - E_0 r\cos\theta.\end{aligned} \tag{5–18}$$

Note that the field **E** inside the sphere is uniform, but is smaller than the field at infinity by the ratio $3/(\kappa+2)$. It is also seen that the induced field of the sphere in the region outside the sphere is that of a dipole whose moment is

$$\mathbf{p} = 4\pi\epsilon_0 a^3\left(\frac{\kappa-1}{\kappa+2}\right)\mathbf{E}_0. \tag{5–19}$$

Let a quantity L be what is known as the depolarization factor for a dielectric body, defined as

$$L = \frac{|\mathbf{E}_0| - |\mathbf{E}_{\text{inside}}|}{\epsilon_0|\mathbf{P}_{\text{inside}}|}. \tag{5–20}$$

It will be remembered that $\mathbf{P} = \epsilon_0(\kappa-1)\mathbf{E}$, from Eq. (2–12). Then for a sphere $L = \frac{1}{3}$, while for a thin rod oriented parallel to the field $L \simeq 0$, and for a thin disk oriented normal to the field $L = 1$. Thus the electric field within a dielectric body in a uniform field is always smaller than the field at a large distance, while the dielectric displacement is always larger.

5-5 The potential of an arbitrary axially-symmetric spherical potential distribution. The potential at any point in space corresponding to a given potential distribution, $\phi(a,\theta)$, over a sphere of radius a, can also be written in terms of Legendre polynomials. From Eqs. (5-5) and (5-9), we shall have two potentials, one valid inside and one valid outside the surface:

$$\begin{aligned} \underset{r<a}{\phi} &= \sum_{n=0}^{\infty} A_n r^n P_n(\mu), \\ \underset{r>a}{\phi} &= \sum_{n=0}^{\infty} B_n r^{-n-1} P_n(\mu). \end{aligned} \tag{5-21}$$

We can determine A_n and B_n by taking advantage of the orthogonality of the functions $P_n(\mu)$ if we know the normalizing factor, i.e., the integral of P_n^2 over the range of its variable. Stated in a general way, so as to include orthogonality, the integral needed is

$$\int_{-1}^{+1} P_n(\mu) P_m(\mu)\, d\mu = \frac{2}{2n+1}\,\delta_{mn}, \tag{5-22}$$

where $\delta_{mn} = 1$ if $m = n$, and is zero otherwise.

Either of Eqs. (5-21) may then be equated to $\phi(a,\theta)$ for $r = a$, the resulting equation multiplied by $P_m(\mu)$ and integrated from $\mu = -1$ to $\mu = +1$; only one term of the series survives, namely, that for which $n = m$. As a result of solving for A_n,

$$A_n = \frac{2n+1}{2a^n} \int_{-1}^{+1} \phi(a,\theta) P_n(\mu)\, d\mu \tag{5-23}$$

and

$$B_n = a^{2n+1} A_n. \tag{5-24}$$

The potentials therefore become

$$\begin{aligned} \underset{r<a}{\phi} &= \sum_{n=0}^{\infty} \frac{2n+1}{2a^n} r^n P_n(\mu) \int_{-1}^{+1} \phi(a,\theta') P_n(\mu')\, d\mu', \\ \underset{r>a}{\phi} &= \sum_{n=0}^{\infty} \frac{2n+1}{2} a^{n+1} r^{-n-1} P_n(\mu) \int_{-1}^{+1} \phi(a,\theta') P_n(\mu')\, d\mu'. \end{aligned} \tag{5-25}$$

5-6 The potential of a charged ring. Let us consider the potential of a charged ring, of total charge q, located at a distance r_0 from the origin, as seen in Fig. 5–4, with r_0 making an angle θ_0 with the axis of symmetry. If the distance from the origin along the polar axis is called z, the potential along the polar axis, found by expanding $1/R$ in the Coulomb potential by Eq. (5–13), is

$$\underset{r<r_0}{\phi}(z,0) = \frac{q}{4\pi\epsilon_0 r_0}\sum_{n=0}^{\infty}\left(\frac{z}{r_0}\right)^n P_n(\cos\theta_0),$$

$$\underset{r>r_0}{\phi}(z,0) = \frac{q}{4\pi\epsilon_0 r_0}\sum_{n=0}^{\infty}\left(\frac{r_0}{z}\right)^{n+1} P_n(\cos\theta_0). \tag{5–26}$$

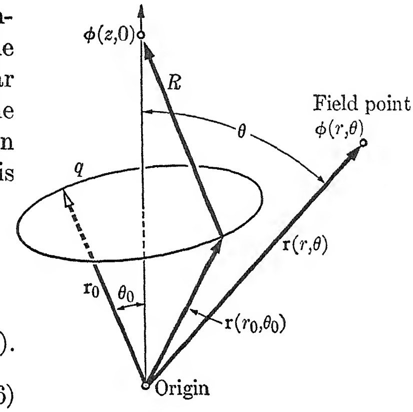


FIG. 5–4 Coordinates for finding the potential due to a charged ring.

It is evident from the series of Eq. (5–12) that $P_n(1) = 1$. Therefore the potential at a general point, not lying on the polar axis, may be found by multiplying the nth term in the series by $P_n(\mu)$ and writing r for z:

$$\begin{aligned}\underset{r<r_0}{\phi}(r,\theta) &= \frac{q}{4\pi\epsilon_0 r_0}\sum_{n=0}^{\infty}\left(\frac{r}{r_0}\right)^n P_n(\cos\theta_0)P_n(\cos\theta),\\ \underset{r>r_0}{\phi}(r,\theta) &= \frac{q}{4\pi\epsilon_0 r_0}\sum_{n=0}^{\infty}\left(\frac{r_0}{r}\right)^{n+1} P_n(\cos\theta_0)P_n(\cos\theta).\end{aligned} \tag{5–27}$$

The uniqueness theorem is essential to justify the argument that led to the above result.

5–7 Problems not having axial symmetry. If the geometry of a problem does not have axial symmetry, but does have spherical boundaries, the potential solutions for the problem may be written down formally just as in the examples above. The constant m must be allowed to have values different from zero in order to represent the asymmetry, the spherical harmonics $Y_n^m(\theta,\varphi)$ are characterized by two parameters, and the series solution is a double sum: for each n there is a set of values for m, with the condition that $m \leq n$. The functions Φ are simply the Fourier functions, and the functions Θ are $P_n^m(\mu)$ and $Q_n^m(\mu)$, the associated Legendre functions. Thus the form of the separated functions is different, but no principles are involved which have not been exemplified in problems for which Φ is constant. We shall therefore omit a detailed discussion of such problems. We shall also omit the consideration of problems with conical boundaries that exclude the polar axis from the range of potentials, and which require the use of $Q_n^m(\mu)$ in order that the boundary conditions be satisfied. Abundant references for the treatment of many such problems are listed at the end of the chapter.

5–8 The solution of Laplace's equation in cylindrical coordinates. Laplace's equation in the cylindrical coordinates r, φ, z is

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\phi}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2\phi}{\partial\varphi^2}+\frac{\partial^2\phi}{\partial z^2}=0. \tag{5–28}$$

Separating by means of the product functions,

$$\phi(r,\varphi,z)=R(r)\Phi(\varphi)Z(z), \tag{5–29}$$

we obtain

$$r\frac{d}{dr}\left(r\frac{dR}{dr}\right)+(k^2r^2-n^2)R=0, \tag{5–30}$$

$$\frac{d^2\Phi}{d\varphi^2}+n^2\Phi=0, \tag{5–31}$$

$$\frac{d^2Z}{dz^2}-k^2Z=0, \tag{5–32}$$

where k and n are the separation parameters. Equation (5–30) is known as Bessel's equation, and its solutions are called Bessel functions. The character of the solution will depend markedly on the sign of the separation constant, i.e., on whether n and k are real or imaginary. If solutions are desired which are single valued in the azimuth angle φ, then the solutions must be periodic in φ and n must be a real integer. If k is real, the solution $Z(z)$ is a real exponential, and the radial solutions are combinations of the Bessel functions designated by $J_n(kr)$ and $N_n(kr)$. Excellent treatments of Bessel functions and their properties are listed at the end of the chapter. We need remark here only that J_n and N_n are oscillatory functions of their arguments, and that J_n is the solution which is regular at $r=0$, a point for which Eq. (5–30) has a singularity. For real n and k, then, the integrals are of the form:

$$\begin{aligned}
R(r)&=A_nJ_n(kr)+B_nN_n(kr), && k\neq 0,\\
R(r)&=Ar^n+Br^{-n}, && k=0,\\
\Phi(\varphi)&=C_n\cos n\varphi+D_n\sin n\varphi, && n\neq 0,\\
\Phi(\varphi)&=C\varphi+D, && n=0,\\
Z(z)&=E_ke^{kz}+F_ke^{-kz}, && k\neq 0,\\
Z(z)&=Ez+F, && k=0.
\end{aligned} \tag{5–33}$$

If n and k are both zero,

$$\phi = (A \ln r + B)(C\varphi + D)(Ez + F).$$

We shall illustrate the use of these cylindrical functions by outlining the solution for a sample problem. Consider a case which has azimuthal symmetry, so that $n = 0$. If the solution extends to infinity radially, so that there are no cylindrical boundaries that might impose restrictions on the radial function $R(r)$, there are correspondingly no restrictions on k and the solution involves an integral over all k. The integral will have the form

$$\int_0^\infty e^{\pm kz} f(k) J_0(kr)\, dk. \tag{5-34}$$

Instead of determining a set of coefficients in a sum, we have now to determine the value of the function $f(k)$. The potential of a point charge may be expressed by an integral of this kind and, in fact, is given in terms of a mathematical identity

$$\frac{1}{R} = \frac{1}{\sqrt{r^2 + z^2}} = \int_0^\infty e^{\pm kz} J_0(kr)\, dk, \tag{5-35}$$

where the exponential is positive for $z < 0$ and negative for $z > 0$. The Coulomb potential $\phi = q/4\pi\epsilon_0 R$ is

$$\phi = \frac{q}{4\pi\epsilon_0} \int_0^\infty e^{\pm kz} J_0(kr)\, dk. \tag{5-36}$$

The potential of Eq. (5-36) can then be used in combination with the induced potential of the form Eq. (5-34) to write down the solution of a problem corresponding to plane boundaries normal to the z-axis and under the influence of a point charge located at the origin. This layer structure, shown in Fig. 5-5, composed of several layers of different inductive capaci-

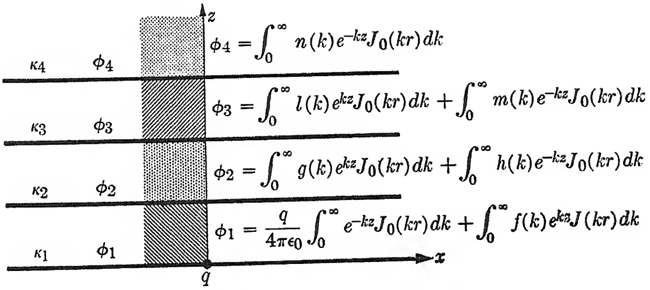

z
$\phi_4 = \int_0^\infty n(k) e^{-kz} J_0(kr)\, dk$
κ_4
ϕ_4
$\phi_3 = \int_0^\infty l(k) e^{kz} J_0(kr)\, dk + \int_0^\infty m(k) e^{-kz} J_0(kr)\, dk$
κ_3
ϕ_3
$\phi_2 = \int_0^\infty g(k) e^{kz} J_0(kr)\, dk + \int_0^\infty h(k) e^{-kz} J_0(kr)\, dk$
κ_2
ϕ_2
$\phi_1 = \frac{q}{4\pi\epsilon_0} \int_0^\infty e^{-kz} J_0(kr)\, dk + \int_0^\infty f(k) e^{kz} J(kr)\, dk$
κ_1
ϕ_1
x
q

FIG. 5-5 Parallel layers of dielectric materials under the influence of a point charge q.

ties, κ_1, κ_2, etc., has the potentials shown in the figure. If we apply the boundary conditions at all of the interfaces and equate the functions under the integral sign, there will be a sufficient number of equations to determine the functions of k, and therefore the solution in every layer.

If the cylindrical solution is required to be periodic in the z-direction, k must be imaginary, and the solutions of the radial equation are Bessel functions of an imaginary variable which are usually designated by I_n and K_n. Just as in rectangular and spherical coordinates, we note that it is impossible for all three factors to be oscillatory functions of their respective variables.

5–9 Application of cylindrical solutions to potential problems. Lists of the mathematical properties of Bessel functions necessary for the solution of actual physical problems are found in the references. To solve problems involving conducting cylindrical boundaries, we must make use of tables, much as we make use of trigonometric tables for circular functions. Inside a grounded cylindrical conductor of radius r_0, for example, each solution must satisfy the condition $J_n(kr_0) = 0$, so that k is limited to a discrete set of values k_l, where l is an index numbering the zeros of J_n from 1 to infinity. The entire solution is, in general, a double sum over n and l, with coefficients determined so as to satisfy the boundary conditions. The orthogonality of the functions follows from the general proof of Section 3–6, or it may be shown from the differential equation (5–30) that $\int_0^{r_0} J_n(k_l r)J_n(k_{l'}r)r\,dr = 0$ if $l \neq l'$. The orthogonality of the partial potentials corresponding to different n is obvious from the form of Φ.

The method can be adequately illustrated even with loss of complete generality by the problem of a point charge at the origin of coordinates within a grounded conducting cylinder about the z-axis. The simplification here arises from the azimuthal symmetry, so that the entire potential is independent of φ, and the radial solutions are confined to $J_0(k_l r)$. The desired solution has the form

$$\phi_1 = \sum_{l=1}^{\infty} A_l e^{k_l z} J_0(k_l r), \qquad \text{for } z < 0,$$

$$(5\text{–}37)$$

$$\phi_2 = \sum_{l=1}^{\infty} A_l e^{-k_l z} J_0(k_l r), \qquad \text{for } z > 0,$$

where k_l is such that at the radius of the cylinder $r = r_0$, $J_0(k_l r_0) = 0$.

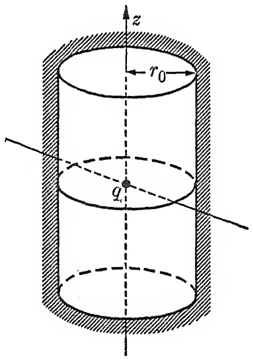


FIG. 5–6 Point charge inside a grounded conducting cylinder.

The flux condition at the plane $z = 0$ can be stated by means of a two-dimensional δ-function defined by the relation

$$2\pi \int_0^{r_0} \delta(r) r\, dr = 1, \tag{5–38}$$

so that

$$\begin{aligned} \frac{q\, \delta(r)}{\epsilon_0} &= \left[\frac{\partial \phi_1}{\partial z} - \frac{\partial \phi_2}{\partial z} \right]_{z=0} \\ &= 2\Sigma k_l A_l J_0(k_l r). \end{aligned} \tag{5–39}$$

If we multiply both sides of Eq. (5–39) by $J_0(k_{l'} r) r\, dr$ and integrate over r, we obtain

$$\frac{q J_0(0)}{2\pi\epsilon_0} = 2k_{l'} A_{l'} \int_0^{r_0} [J_0(k_{l'} r)]^2 r\, dr,$$

or

$$A_l = \frac{q/\epsilon_0}{4\pi k_l} \frac{J_0(0)}{\displaystyle\int_0^{r_0} [J_0(k_l r)]^2 r\, dr}. \tag{5–40}$$

Now the Bessel functions are defined in such a way that $J_0(0) = 1$, and it may be shown by multiplying Eq. (5–30) by $(dJ_n/dr) r^2\, dr$ and integrating by parts that

$$\int_0^{r_0} [J_0(k_l r)]^2 r\, dr = \frac{r_0^2}{2} [J_1(k_l r_0)]^2$$

if we take account of the boundary condition at r_0 and the mathematical identity

$$\frac{dJ_n}{d(kr)} = -J_{n+1} + \frac{n}{kr} J_n. \tag{5–41}$$

Therefore

$$A_l = \frac{q/\epsilon_0}{2\pi k_l} \frac{1}{r_0^2 [J_1(k_l r_0)]^2} \tag{5–42}$$

and

$$\begin{aligned} \phi_1 &= \frac{q}{2\pi r_0^2 \epsilon_0} \sum_{l=1}^{\infty} e^{k_l z} \frac{J_0(k_l r)}{k_l [J_1(k_l r_0)]^2}, \\ \phi_2 &= \frac{q}{2\pi r_0^2 \epsilon_0} \sum_{l=1}^{\infty} e^{-k_l z} \frac{J_0(k_l r)}{k_l [J_1(k_l r_0)]^2}. \end{aligned} \tag{5–43}$$

A generalization of this procedure can be used to determine the Green's function for the interior of a conducting cylinder if the point charge is located at an arbitrary position $r = a$, $\varphi = \alpha$, $z = z_0$. In that case the potentials are written separately for $z < z_0$, $z > z_0$, and the flux condition is expressed in terms of the generalized two-dimensional δ-function for which

$$\int \delta(r - a)\, \delta(\varphi - \alpha) r\, dr\, d\varphi = 1. \tag{5-44}$$

The resulting Green's function is a double sum over n and l which reduces to Eq. (5-43) with $q = 1$ for the unit charge at the origin.

SUGGESTED REFERENCES

Of the references listed in earlier chapters those by Smythe, Jeans, Stratton, Weber, and Morse and Feshbach will probably be most useful for solving more elaborate potential problems than we have included. The properties of Legendre and Bessel functions are developed in an elementary way by Byerly in *Fourier's Series and Spherical Harmonics*, and more complete treatments are to be found in:

T. M. MacRobert, *Spherical Harmonics.*
G. N. Watson, *Theory of Bessel Functions.*

EXERCISES

1. What is the potential distribution inside a spherical region bounded by two conducting hemispheres at potentials $\pm V/2$, respectively?

2. Find the potential at points outside a spherical volume distribution of charge in which the charge density is proportional to the distance from a diametral plane.

3. The equation of the surface of a conductor is $r = a[1 + \delta P_n(\cos\theta)]$, where $\delta \ll 1$. Show that if the conductor is placed in a uniform field $\mathbf{E}$ parallel to the polar axis the surface charge density is given by

$$\sigma = \sigma_0 + \delta \left\{ \frac{3n\epsilon_0 E}{2n+1} [(n+1)P_{n+1}(\cos\theta) + (n-2)P_{n-1}(\cos\theta)] \right\},$$

where σ_0 is the induced charge density for $\delta = 0$.

4. A uniform field E_m is set up in a medium of specific inductive capacity κ. Prove that the field inside a spherical cavity in the medium is given by

$$E = \frac{3\kappa E_m}{2\kappa + 1}.$$

5. A conducting sphere of radius a carrying charge q is placed in a uniform field of strength E_0. Find the potential everywhere. What is the dipole moment of the induced charge on the sphere?

6. What is the capacity of a thin hemispherical shell? (*Hint*: Invert a charged circular disk. The solution of this problem was originally due to Kelvin.)

7. A solid dielectric sphere of radius a has a sector removed so that it fits the edge of an infinite conducting wedge of external angle α, its surface meeting the

wedge faces orthogonally. If the wedge is charged, show that the potentials inside and outside the sphere are

$$\phi_{\text{in}} = \text{const.}\,\frac{\alpha + 2\pi}{\alpha(1+\kappa)}\,(r\sin\theta)^{\pi/\alpha}\cos\frac{\pi\varphi}{\alpha},$$

$$\phi_{\text{out}} = \text{const.}\left[r^{\pi/\alpha} - \frac{\pi(\kappa-1)a^{\pi/\alpha}}{\alpha+\pi(\kappa+1)}\left(\frac{a}{r}\right)^{(\pi/\alpha)+1}\right](\sin\theta)^{\pi/\alpha}\cos\frac{\pi\varphi}{\alpha}.$$

(*Hint*: Note that the geometry of the problem permits only a single surface harmonic. Which is it?) (Smythe)

8. Show that the potential due to a disk of radius a carrying a charge q is

$$\phi(r,z) = \frac{q}{4\pi\epsilon_0 a}\int_0^\infty e^{-k|z|}J_0(kr)\,\frac{\sin ka}{k}\,dk$$

in cylindrical coordinates. Can you suggest other methods for the solution of this problem?

CHAPTER 6

ENERGY RELATIONS AND FORCES IN THE ELECTROSTATIC FIELD

Our discussion of the electrostatic field has thus far been based entirely on a single experimental law, namely Coulomb's law of Eq. (1–24) for the action-at-a-distance force between two point charges. The electric field has been introduced as an intermediate agent whose purpose is to simplify the description of the interaction between charges. The question of the reality of the field as an independent physical entity therefore does not arise in these considerations. Maxwell attempted to ascribe a larger degree of physical reality in a direct mechanical sense to the electric field than will be necessary for our purpose, since the fundamental reason for attributing physical reality to the field will not become apparent until nonstatic effects are discussed. However, if the field is a truly valid representation of the experimental facts it is necessary that all of the mechanical properties of an electrically interacting system can be described either in terms of the sources which partake of the interaction or in terms of the fields themselves which are produced by the sources.

This means that the detailed nature of the sources should not influence the action of a field on a given system of charges. The description of the electric field alone must be a sufficient description to determine what interaction occurs if a number of charges are introduced at given points in the field. This interaction must be formulated in terms of the field itself, and not depend explicitly on the configuration of the charges that produce the field.

It should therefore be possible to develop a field theory in which we can describe the over-all mechanical properties, such as energy and force, equivalently in terms of the charges which are the sources of the field, or in terms of integrals over the field produced by the charges. The only criterion for the correctness of such over-all relations when expressed in terms of the electrostatic field shall be that the results are equivalent to those which are obtained from a direct consideration of the action-at-a-distance interaction of the charges responsible for the field.

6–1 Field energy in free space. Let us consider a set of charges q_i in free space. The work done in the course of physical assembly of these already-created charges, which are initially located an infinite distance apart, is given by

$$W = \tfrac{1}{2} \sum_{i=1}^{n} q_i\phi_i, \tag{6–1}$$

where ϕ_i is the potential produced at the position of q_i by all the other charges. By assembling these charges we have changed the energy of the system and, since all of the forces are conservative, we can identify this expression for the work of assembly with the energy of the system. This energy must be stored somewhere. Just as in mechanics, however, the location that one selects as the place of energy storage depends on one's point of view.

If, for example, we consider two masses at the ends of a compressed spring, we have a system that possesses potential energy which will be released if the spring is allowed to expand. In the expansion the masses will acquire kinetic energy. The physical location of the energy in this mechanical system in its initial condition is not necessarily in the spring. Phenomenologically the masses may be considered to be initially in regions of higher potential energy than they are after the expansion of the spring. Equation (6–1) corresponds to the latter point of view. We shall now try to transform Eq. (6–1) to an expression which would make it appear as if the electrical energy resides in the (figuratively) "elastic" quality of the electric field, as would be required in order to correspond to the point of view that the energy of the mass-spring system resides in the spring.

The expression obtained by Maxwell for the energy in an electric field, expressed as a volume integral over the field, is

$$U = \frac{\epsilon_0}{2} \int E^2 \, dv. \tag{6–2}$$

The integral extends over all space. We shall now show that the field energy U is, in fact, the same as the assembly work W. To show this, let us introduce at any arbitrary field point the partial fields $\mathbf{E}_i$, each being the Coulomb field of only one of the point charges q_i. $\mathbf{E}$ and E^2 are then given by

$$\begin{aligned} \mathbf{E} &= \sum_{i=1}^{n} \mathbf{E}_i, \\ E^2 &= \sum_{i=1}^{n} E_i^2 + \sum_{i=1}^{n} \sum_{j=1}^{n}{}' \, \mathbf{E}_i \cdot \mathbf{E}_j, \end{aligned} \tag{6–3}$$

where the prime on the summation symbol indicates that the term for which $i = j$ is omitted, since such terms are grouped separately in the first summation. For point charges, the first term in the summation makes an infinite contribution to the integral over E^2 in Eq. (6–2). However, this infinite term is independent of the relative position of the charges and therefore it must represent the work necessary to create the charges from

AGRICULTURAL

an arbitrary zero point of energy. We shall therefore designate

$$U_s = \frac{\epsilon_0}{2} \int \sum_{i=1}^{n} E_i^2 \, dv \tag{6–4}$$

as the self-energy of the system. The introduction of a finite radius for the elementary charges enables us to avoid infinite self-energy so long as the charges are stationary, as indicated in a problem at the end of this chapter. Later we shall find that a finite radius involves difficulties when moving charges are considered.

With Eq. (6–4) the Maxwell field energy expression, Eq. (6–2), then becomes

$$U = U_s + \frac{\epsilon_0}{2} \sum_{i=1}^{n} \int \mathbf{E}_i \cdot (-\nabla \Sigma_j' \phi_j) \, dv, \tag{6–5}$$

where $\Sigma_j' \phi_j$ denotes the potential at any point due to all of the charges except the ith charge. Using the vector identity

$$\nabla \cdot (\mathbf{A}\phi) = \phi \nabla \cdot \mathbf{A} + \mathbf{A} \cdot \nabla \phi \tag{6–6}$$

in order to perform an integration by parts, we obtain

$$U = U_s - \frac{\epsilon_0}{2} \sum_{i=1}^{n} \int [\nabla \cdot (\mathbf{E}_i \Sigma_j' \phi_j) - \Sigma_j' \phi_j \nabla \cdot \mathbf{E}_i] \, dv. \tag{6–7}$$

Except at the position of the ith charge, $\nabla \cdot \mathbf{E}_i$ vanishes everywhere. Hence we may write $\nabla \cdot \mathbf{E}_i = q_i \, \delta(r_i)/\epsilon_0$, and hence

$$U = U_s - \frac{\epsilon_0}{2} \sum_{i=1}^{n} \int \mathbf{E}_i (\Sigma_j' \phi_j) \cdot d\mathbf{S} + \frac{\epsilon_0}{2} \sum_{i=1}^{n} \phi_i^0 \frac{q_i}{\epsilon_0}, \tag{6–8}$$

where ϕ_i^0 is $\Sigma_j' \phi_j$ evaluated at the position of the ith charge, i.e., the potential at q_i due to the other charges. The surface term can be made arbitrarily small by letting the boundary surface go to infinity, since the fields decrease at least as the inverse second power, the potential at least as the inverse first power, and the differential area of integration increases only as the square of the distance. If we consider that the integral in Eq. (6–2) extends over all of space where there is a field, then this integral, as a result of Eq. (6–8), reduces to

$$U = U_s + \tfrac{1}{2} \sum_{i=1}^{n} \phi_i^0 q_i. \tag{6–9}$$

The second term of this final equation is identical with the expression for the work of assembly of the charges from infinity, while the first term U_s is the self-energy corresponding to the energy used in the creation of the charges themselves. The analysis shows that Eqs. (6–2) and (6–9) correspond to the same energy, but Eq. (6–2) expresses this energy as a volume integral of an energy density $\epsilon_0 E^2/2$ extending over all of space. It is not possible to ascertain experimentally whether the energy resides in the field or is possessed by the charges which produce the field.

6–2 Energy density within a dielectric. In the presence of a dielectric Eq. (6–2) can under certain conditions be replaced by

$$U = \tfrac{1}{2}\int \mathbf{E}\cdot\mathbf{D}\,dv. \tag{6–10}$$

To show this, let us consider the change of energy when a small increment of true charge $\delta\rho$ is added to the field. We shall assume rigid boundaries, or rigid constraints on the medium, so that no work is done on mechanical constraints. In the case of continuous charges the self-energy problem disappears, and the work done is given by

$$\delta W = \int \phi\,\delta\rho\,dv = \int \phi\,\delta(\nabla\cdot\mathbf{D})\,dv = \int \phi\nabla\cdot\delta\mathbf{D}\,dv. \tag{6–11}$$

The vector theorem of Eq. (6–6) and Gauss's theorem enable us to write

$$\begin{aligned}\delta W &= \int \nabla\cdot(\phi\delta\mathbf{D})\,dv - \int \delta\mathbf{D}\cdot\nabla\phi\,dv\\ &= \int (\phi\delta\mathbf{D})\cdot d\mathbf{S} - \int \delta\mathbf{D}\cdot\nabla\phi\,dv.\end{aligned} \tag{6–12}$$

If we drop the surface term, as we did in the derivation of Eq. (6–9), we obtain

$$\delta W = -\int \delta\mathbf{D}\cdot\nabla\phi\,dv = \int \mathbf{E}\cdot\delta\mathbf{D}\,dv. \tag{6–13}$$

This increment of work usually cannot be integrated unless **E** is a given function of **D**. If, for example, **E** and **D** are related by a dielectric constant κ, which may be a function of position but not of **E**, then the energy resulting from the integration of the work increment from $\mathbf{D} = 0$ to $\mathbf{D} = \mathbf{D}$ is

$$\begin{aligned}U &= \int_0^D \delta W = \int_0^D \int \mathbf{E}\cdot\delta\mathbf{D}\,dv = \iint_0^E \frac{\kappa\epsilon_0\,\delta(E^2)}{2}\,dv\\ &= \tfrac{1}{2}\int \kappa\epsilon_0 E^2\,dv = \tfrac{1}{2}\int \mathbf{E}\cdot\mathbf{D}\,dv.\end{aligned} \tag{6–14}$$

At least for the simple case, therefore, in which **E** and **D** are proportional, Eq. (6–10) is justified.

6–3 Thermodynamic interpretation of *U*. The assumption of a dielectric constant κ that does not change with time and is a function only of position implies that the process of change of field is an isothermal process, since the dielectric constant is usually a function of the temperature. If energy enters the dielectric it may produce heat and cause a variation of κ with time. To assure isothermal behavior, the dielectric material in question must be in contact with a heat bath which can abstract heat from it to maintain a constant temperature. Thus we cannot equate the increment of work done, δW of Eq. (6–13), to the increase in total energy, since heat changes are also involved. Equation (6–13) does, however, represent the maximum work which can be extracted at a later time from the total electrical field energy.

Thermodynamically, the maximum work which can be obtained from a system under isothermal conditions is the free energy F of the system, not the total energy. This means that in the presence of dielectrics the expression $U = \frac{1}{2}\int \mathbf{E} \cdot \mathbf{D}\, dv$ cannot be identified with the total energy of the system, but can only be identified with the thermodynamic free energy. Of course, the distinction vanishes when no materials with temperature-dependent dielectric properties are present in the field. In the thermodynamic sense, the electric field $\mathbf{E}$ is analogous to gas pressure and the displacement $\mathbf{D}$ is analogous to volume.

The correct expression for the total energy can be easily derived. If we call the total energy U and the entropy S, and let T be the absolute temperature, then,

$$F = U - TS, \quad dF = dU - T\,dS - S\,dT. \tag{6–15}$$

We have identified $\int \mathbf{E} \cdot d\mathbf{D}\, dv$ as the increment of work at constant temperature,

$$dF|_T = dU - T\,dS = \int \mathbf{E} \cdot d\mathbf{D}\, dv.$$

Hence, from Eq. (6–15),

$$S = -\left.\frac{\partial F}{\partial T}\right|_D = \int \frac{\epsilon_0}{2} E^2 \frac{d\kappa}{dT}\, dv = \int \frac{\mathbf{E} \cdot \mathbf{D}}{2} \frac{1}{\kappa} \frac{d\kappa}{dT}\, dv,$$

and

$$U = F + TS = \int \frac{\epsilon_0}{2} E^2 \frac{d(T\kappa)}{dT}\, dv = \int \frac{\mathbf{E} \cdot \mathbf{D}}{2} \frac{1}{\kappa} \frac{d(T\kappa)}{dT}\, dv. \tag{6–16}$$

The heat absorbed during application of the field is thus

$$\delta Q = T\,dS = \int \mathbf{E} \cdot d\mathbf{D} \frac{T}{\kappa} \frac{d\kappa}{dT}\, dv. \tag{6–17}$$

If, for example, the specific inductive capacity has the form

$$\kappa = 1 + \chi = 1 + A/T \tag{6-18}$$

with A constant, as in the case of gases composed of molecules having permanent dipole moments, then $d\kappa/dT$ is less than zero, and hence $\delta Q < 0$ if $\delta D > 0$. Hence heat will be given off when the field is applied, and conversely.

6-4 Thomson's theorem. From now on we shall use only the free-energy density in our considerations, although we shall follow the conventions of electromagnetic nomenclature and designate it by U instead of F. This will enable us to equate changes in the free energy directly to the mechanical work quantities responsible for them, without making it necessary to include thermal quantities in the energy balance. The free-energy expression of Eq. (6-10), applicable even in the presence of dielectrics, behaves in electrical problems in the same manner as does the chemical free energy in chemical kinetics: a reaction will proceed until the free energy reaches a minimum value. In the electric case, charges on a conductor will redistribute themselves in such a way that the entire free-field energy will be minimized.

We can show this directly. Let us consider a virtual process in which charges in equilibrium on a conductor are displaced by an infinitesimal amount along the constant potential conductor surfaces in such a way that the total charge remains unchanged. The variation of free energy is given by

$$\delta U = \tfrac{1}{2}\int \kappa\epsilon_0\, \delta(E^2)\, dv = \int \mathbf{E}\cdot\delta\mathbf{D}\, dv. \tag{6-19}$$

If we apply Gauss's theorem and let the surface term vanish, we obtain

$$\delta U = \int \mathbf{E}\cdot\delta\mathbf{D}\, dv = \int (-\nabla\phi\cdot\delta\mathbf{D})\, dv = \int [\phi\nabla\cdot\delta\mathbf{D} - \nabla\cdot\delta(\mathbf{D}\phi)]\, dv$$

$$\delta U = \Sigma\phi_i \int \delta\rho_i\, dv - \int \delta(\mathbf{D}\phi)\cdot d\mathbf{S} = \Sigma\phi_i\int \delta\rho_i\, dv = 0. \tag{6-20}$$

The summation extends over each individual conductor which, since it is at equilibrium, is at constant potential ϕ_i. The final expression vanishes because the total charge on each conductor is unchanged, and thus the variation of the free energy is zero when a system is in equilibrium. This theorem, usually known as Thomson's theorem, shows that the free energy is actually an extremum at equilibrium.

The term

$$U_v = \frac{\mathbf{E} \cdot \mathbf{D}}{2} \tag{6-21}$$

is known as the energy density (more accurately, of course, as the free-energy density) of the electrostatic field. It is a density in the sense that its volume integral gives the over-all energy of the field. On the other hand, in the same sense as it was impossible to localize the energy either in the field or in the source charges, it is also impossible to associate energy in a definite way with each specific volume of field in a manner which can be verified by experiment.

In deriving the energy expression it is assumed that the medium is held at rest and hence no work is done in motion against forces. This implies that the virtual process of assembling the charges in the dielectrics is a process with particular constraints. The resultant energy expression is nevertheless general, since no nonconservative forces are involved. We shall next consider the more general virtual process permitting mass motion; Eq. (6–10) will still be applicable, however, since the final field energy is independent of history.

6–5 Volume forces in the electrostatic field. The force per unit volume that acts on a dielectric body when it is under the influence of an external electrostatic field may be derived from the energy principle. The variation in free energy δU when a unit volume of the dielectric undergoes a virtual displacement $\delta\mathbf{x}$ is given in terms of $\mathbf{F}_v$, the force per unit volume, by*

$$\delta U = -\int \mathbf{F}_v \cdot \delta\mathbf{x}\, dv. \tag{6-22}$$

In view of the fact that the magnitude of the virtual displacement $\delta\mathbf{x}$ is quite arbitrary, we can identify $\mathbf{F}_v$ in Eq. (6–22) with the true volume force. This information may be put in a different way: if $\mathbf{u}$ is an arbitrary velocity field within a dielectric, then the rate at which energy is lost by the field is given by

$$\frac{dU}{dt} = -\int \mathbf{F}_v \cdot \mathbf{u}\, dv, \tag{6-23}$$

where $\mathbf{F}_v$ again represents the volume force.

Let us now consider the energy change due to both a change $\delta\rho$ in true charge distribution and a change $\delta\kappa$ in specific inductive capacity caused by the displacements. From Eq. (6–21),

---

* It is assumed here that the virtual velocities $\mathbf{u}$ corresponding to the virtual displacements $\delta\mathbf{x}$ are sufficiently slow so that the process is both reversible and isothermal. Under these conditions the change in free energy can be equated to the mechanical work done.

$$\delta U = \tfrac{1}{2}\delta \int \mathbf{E} \cdot \mathbf{D}\, dv$$

$$= \frac{1}{2\epsilon_0} \int D^2\, \delta(1/\kappa)\, dv + \int \mathbf{E} \cdot \delta\mathbf{D}\, dv \tag{6–24}$$

$$= -\frac{\epsilon_0}{2} \int E^2\, \delta\kappa\, dv + \int \mathbf{E} \cdot \delta\mathbf{D}\, dv. \tag{6–25}$$

Again making use of integration by parts, we see that

$$\int \mathbf{E} \cdot \delta\mathbf{D}\, dv = -\int \nabla\phi \cdot \delta\mathbf{D}\, dv \tag{6–26}$$

$$= \int \phi \nabla \cdot (\delta\mathbf{D})\, dv = \int \phi \delta \nabla \cdot \mathbf{D}\, dv$$

$$= \int \phi\, \delta\rho\, dv. \tag{6–27}$$

Hence

$$\frac{dU}{dt} = \int \left(\phi \frac{\partial \rho}{\partial t} - \frac{\epsilon_0}{2} E^2 \frac{\partial \kappa}{\partial t} \right) dv. \tag{6–28}$$

To arrive at an expression for $\mathbf{F}_v$, we must express the time derivatives $\partial\rho/\partial t$ and $\partial\kappa/\partial t$ of Eq. (6–28) in terms of the arbitrary velocity field $\mathbf{u}$. This can be done by means of the hydrodynamic equations of continuity,

$$\nabla \cdot (\rho \mathbf{u}) + \frac{\partial \rho}{\partial t} = 0, \tag{6–29}$$

$$\nabla \cdot (g \mathbf{u}) + \frac{\partial g}{\partial t} = 0, \tag{6–30}$$

which represent respectively the conservation of charge and mass if g is the mass density just as ρ is the charge density. To calculate $\partial\kappa/\partial t$ we must associate the change in dielectric constant with the velocity flow. Since there is net transport of material in a velocity field, the change in dielectric constant can be associated with changes in geometry only if we consider the time history of a volume element that is moving with the velocity $\mathbf{u}$. The total derivative of a particular quantity, such as κ or g, when evaluated so that the observation point for this derivative moves with a chosen volume element in a velocity field, is known as the substantial derivative, and is related to the partial derivatives and to the velocity $\mathbf{u}$ by

$$\frac{D\kappa}{Dt} = \frac{\partial\kappa}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial\kappa}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial\kappa}{\partial z}\frac{\partial z}{\partial t} + \frac{\partial\kappa}{\partial t} = \nabla\kappa\cdot\mathbf{u} + \frac{\partial\kappa}{\partial t}. \tag{6-31}$$

Hence the desired partial derivatives are

$$\begin{aligned} \frac{\partial\kappa}{\partial t} &= -\nabla\kappa\cdot\mathbf{u} + \frac{D\kappa}{Dt}, \\ \frac{\partial g}{\partial t} &= -\nabla g\cdot\mathbf{u} + \frac{Dg}{Dt}. \end{aligned} \tag{6-32}$$

If there is a dielectric equation of state, i.e., a relation which gives the dependence of the dielectric constant on the density, such as the Clausius-Mossotti relation of Eq. (2–39), then the substantial derivative of the dielectric constant can be expressed in terms of the substantial derivative of the density by

$$\frac{D\kappa}{Dt} = \frac{d\kappa}{dg}\frac{Dg}{Dt}. \tag{6-33}$$

The assumption that the dielectric constant depends on the mass density alone includes, of course, the assumption that the virtual processes are isothermal. The factor Dg/Dt of Eq. (6–33) may be evaluated with the aid of Eq. (6–32) and the equation of continuity, so that

$$\frac{D\kappa}{Dt} = \frac{d\kappa}{dg}\left(\frac{\partial g}{\partial t} + \nabla g\cdot\mathbf{u}\right) = \frac{d\kappa}{dg}[\nabla g\cdot\mathbf{u} - \nabla\cdot(g\mathbf{u})] = -\frac{d\kappa}{dg}g\nabla\cdot\mathbf{u}. \tag{6-34}$$

Therefore

$$\frac{\partial\kappa}{\partial t} = -\frac{d\kappa}{dg}g\nabla\cdot\mathbf{u} - \nabla\kappa\cdot\mathbf{u}. \tag{6-35}$$

Equation (6–28) may now be written

$$\frac{dU}{dt} = \int\left[-\phi\nabla\cdot(\rho\mathbf{u}) + \frac{\epsilon_0}{2}E^2\frac{d\kappa}{dg}g\nabla\cdot\mathbf{u} + \frac{\epsilon_0}{2}E^2\nabla\kappa\cdot\mathbf{u}\right]dv. \tag{6-36}$$

The integrand must be brought into the form of the dot product of an expression and the velocity $\mathbf{u}$ if we are to identify the volume force $\mathbf{F}_v$. The first term may be written in this form if we use Eq. (6–6) to integrate by parts and assume that the surface boundary is outside of the dielectric and therefore outside the region of charge density. The result is

$$-\int\phi\nabla\cdot(\rho\mathbf{u})\,dv = \int\rho\nabla\phi\cdot\mathbf{u}\,dv. \tag{6-37}$$

Similarly, the second term can be put into the desired form if we drop a surface term:

$$\frac{\epsilon_0}{2}\int E^2 g\,\frac{d\kappa}{dg}\,\nabla\cdot\mathbf{u}\,dv = -\frac{\epsilon_0}{2}\int\nabla\left(E^2\frac{d\kappa}{dg}\,g\right)\cdot\mathbf{u}\,dv. \tag{6-38}$$

Equation (6–36) therefore becomes

$$\frac{dU}{dt} = \int\left[-\rho\mathbf{E} + \frac{\epsilon_0}{2}\,E^2\nabla\kappa - \frac{\epsilon_0}{2}\,\nabla\left(E^2\frac{d\kappa}{dg}\,g\right)\right]\cdot\mathbf{u}\,dv, \tag{6-39}$$

and by a comparison of Eqs. (6–39) and (6–23) we see that

$$\mathbf{F}_v = \rho\mathbf{E} - \frac{\epsilon_0}{2}\,E^2\nabla\kappa + \frac{\epsilon_0}{2}\,\nabla\left(E^2\frac{d\kappa}{dg}\,g\right) \tag{6-40}$$

is the volume force.

The first term in Eq. (6–40) is the ordinary electrostatic volume force in agreement with Eq. (1–23). The second term represents a force which will appear whenever an inhomogeneous dielectric is in an electric field. The last term, known as the electrostriction term, gives a volume force on a dielectric in an inhomogeneous electric field. Note that the magnitude of the electrostriction term depends explicitly through $d\kappa/dg = (\partial\kappa/\partial g)_T$ on the electrical equation of state of the material. It is interesting to note that this term will never give a net force on a finite region of dielectric if we integrate it over a large enough portion of dielectric so that its extremities are in a field-free region. Under this condition the electrostriction term, being a pure gradient, will integrate out. It is for this reason that the term is frequently omitted, since in the calculation of total forces on dielectric bodies it usually does not contribute. In cases where it can be omitted, however, an incorrect pressure variation within the dielectric is obtained, even though the total force is given correctly.

6–6 Maxwell stress tensor. Before considering any specific examples of the application of Eq. (6–40), we shall reformulate this volume force in terms of its "space stress." In a pure field theory it should be possible to calculate the net force on a given volume element within a dielectric in terms of the field conditions on the surface of the volume element. This implies that the field is the stress-transmitting medium in the same sense that a string tying two weights together is the medium that transmits a force from one weight to the other. This was a point that was emphasized by Maxwell to bring out the importance and the physical reality of field quantities. But, as we have seen before, we can give only an alternate description of the way in which the forces act, and cannot give a definite physical proof of the validity of the field concept exemplified by Eq. (6–40) as compared with the concept of action at a distance. We where b has been set equal to zero for simplicity. Note that additional terms will appear in the stress tensor if the field is not irrotational as assumed above.

The Maxwell tensor is a symmetric tensor of the second rank and can therefore be reduced to three components by transformation to principal axes. The principal values of the matrix can be obtained by solving the secular determinant:

$$|T_{\alpha\beta} - \delta_{\alpha\beta}\lambda| = 0. \tag{6-55}$$

The principal values of the tensor, when $b = 0$, are

$$\lambda_1 = \frac{\kappa\epsilon_0}{2} E^2, \quad \lambda_{2,3} = -\frac{\kappa\epsilon_0}{2} E^2. \tag{6-56}$$

When expressed in terms of principal coordinates the stress tensor therefore takes the simple form:

$$T' = \frac{\kappa\epsilon_0}{2}\begin{pmatrix} E^2 & 0 & 0 \\ 0 & -E^2 & 0 \\ 0 & 0 & -E^2 \end{pmatrix}. \tag{6-57}$$

The principal axes are so oriented that the coordinate axis corresponding to the single root of the secular determinant, λ_1, is parallel to $\mathbf{E}$, while the two axes corresponding to the double roots λ_2 and λ_3 are perpendicular to $\mathbf{E}$. This fact is often expressed qualitatively by stating that the electric field transmits a tension $\kappa\epsilon_0 E^2/2$ parallel to the direction of the field and a contraction of magnitude $\kappa\epsilon_0 E^2/2$ transverse to the direction of the field.

Let us choose a coordinate system in which the x-axis is parallel to the direction of the field, so that $E_y = E_z = 0$, and consider the stress across a surface element, as shown in Fig. 6-1, whose normal makes an angle θ with the x-axis. The stress will then have two components, one parallel to $\mathbf{E}$ and the other perpendicular to $\mathbf{E}$ in the plane defined by $\mathbf{E}$ and the

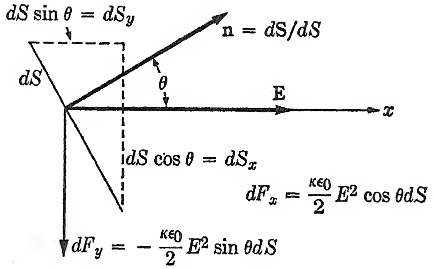

$dS \sin\theta = dS_y$
$\mathbf{n} = d\mathbf{S}/dS$
dS
θ
$\mathbf{E}$
x
$dS \cos\theta = dS_x$
$dF_x = \frac{\kappa\epsilon_0}{2} E^2 \cos\theta dS$
$dF_y = -\frac{\kappa\epsilon_0}{2} E^2 \sin\theta dS$

FIG. 6-1 Geometry for considering stresses at a surface element.

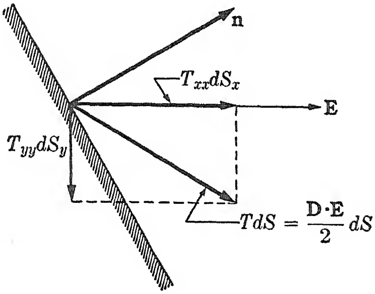


FIG. 6-2 Stresses at a surface.

normal to the surface element. The magnitudes of these stresses are the stress components given by the matrix in Eq. (6-57) multiplied by the surface element components as indicated in the figure. The resultant stress on $d\mathbf{S}$ is the vector sum of the two stress components, as shown in Fig. 6-2. It is seen that the electric field bisects the angle between the normal to the surface and the direction of the resultant stress acting on the surface. This construction is frequently a useful one in the graphical evaluation of the forces on a charged region if an experimental field plot is available or, in the analogous magnetic case to be discussed later, it is useful for the computation of forces on magnetized materials, or on current-carrying conductors.

In the special case of stress transmitted across surfaces either parallel or normal to the electric field, we have the simple situation indicated in the first two parts of Fig. 6-3, where the field transmits a pull of magnitude $ED/2$ across a surface that is normal to the field and a push of magnitude $ED/2$ across a surface that is tangential to the field. A surface that is oriented at 45° to the direction of the field, as also shown in Fig. 6-3, will receive a force that acts parallel to the surface, also of magnitude $ED/2$. These relations can be demonstrated for simple cases such as the attraction and repulsion between two charges of opposite or equal sign.

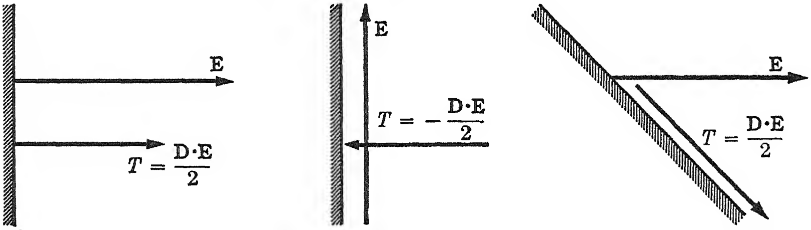

E
$T = \frac{\mathbf{D}\cdot\mathbf{E}}{2}$
E
$T = -\frac{\mathbf{D}\cdot\mathbf{E}}{2}$
E
$T = \frac{\mathbf{D}\cdot\mathbf{E}}{2}$

FIG. 6-3 Stresses for fields perpendicular, parallel, and at an angle of 45° to a surface.

where b has been set equal to zero for simplicity. Note that additional terms will appear in the stress tensor if the field is not irrotational as assumed above.

The Maxwell tensor is a symmetric tensor of the second rank and can therefore be reduced to three components by transformation to principal axes. The principal values of the matrix can be obtained by solving the secular determinant:

$$|T_{\alpha\beta} - \delta_{\alpha\beta}\lambda| = 0. \tag{6–55}$$

The principal values of the tensor, when $b = 0$, are

$$\lambda_1 = \frac{\kappa\epsilon_0}{2} E^2, \quad \lambda_{2,3} = -\frac{\kappa\epsilon_0}{2} E^2. \tag{6–56}$$

When expressed in terms of principal coordinates the stress tensor therefore takes the simple form:

$$T' = \frac{\kappa\epsilon_0}{2}\begin{pmatrix} E^2 & 0 & 0 \\ 0 & -E^2 & 0 \\ 0 & 0 & -E^2 \end{pmatrix}. \tag{6–57}$$

The principal axes are so oriented that the coordinate axis corresponding to the single root of the secular determinant, λ_1, is parallel to **E**, while the two axes corresponding to the double roots λ_2 and λ_3 are perpendicular to **E**. This fact is often expressed qualitatively by stating that the electric field transmits a tension $\kappa\epsilon_0 E^2/2$ parallel to the direction of the field and a contraction of magnitude $\kappa\epsilon_0 E^2/2$ transverse to the direction of the field.

Let us choose a coordinate system in which the x-axis is parallel to the direction of the field, so that $E_y = E_z = 0$, and consider the stress across a surface element, as shown in Fig. 6–1, whose normal makes an angle θ with the x-axis. The stress will then have two components, one parallel to **E** and the other perpendicular to **E** in the plane defined by **E** and the

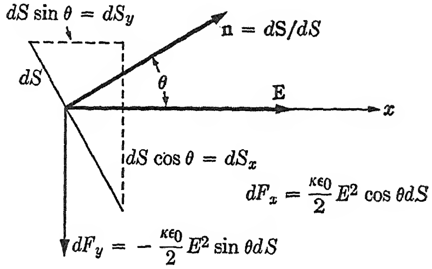

$dS \sin\theta = dS_y$
$\mathbf{n} = d\mathbf{S}/dS$
dS
θ
$\mathbf{E}$
x
$dS \cos\theta = dS_x$
$dF_x = \frac{\kappa\epsilon_0}{2} E^2 \cos\theta dS$
$dF_y = -\frac{\kappa\epsilon_0}{2} E^2 \sin\theta dS$

FIG. 6–1 Geometry for considering stresses at a surface element.

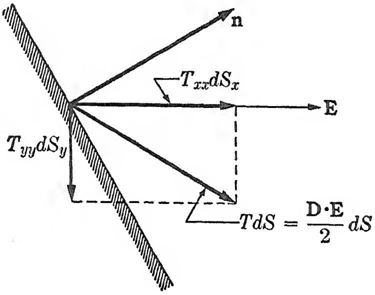


FIG. 6–2 Stresses at a surface.

normal to the surface element. The magnitudes of these stresses are the stress components given by the matrix in Eq. (6–57) multiplied by the surface element components as indicated in the figure. The resultant stress on $d\mathbf{S}$ is the vector sum of the two stress components, as shown in Fig. 6–2. It is seen that the electric field bisects the angle between the normal to the surface and the direction of the resultant stress acting on the surface. This construction is frequently a useful one in the graphical evaluation of the forces on a charged region if an experimental field plot is available or, in the analogous magnetic case to be discussed later, it is useful for the computation of forces on magnetized materials, or on current-carrying conductors.

In the special case of stress transmitted across surfaces either parallel or normal to the electric field, we have the simple situation indicated in the first two parts of Fig. 6–3, where the field transmits a pull of magnitude $ED/2$ across a surface that is normal to the field and a push of magnitude $ED/2$ across a surface that is tangential to the field. A surface that is oriented at 45° to the direction of the field, as also shown in Fig. 6–3, will receive a force that acts parallel to the surface, also of magnitude $ED/2$. These relations can be demonstrated for simple cases such as the attraction and repulsion between two charges of opposite or equal sign.

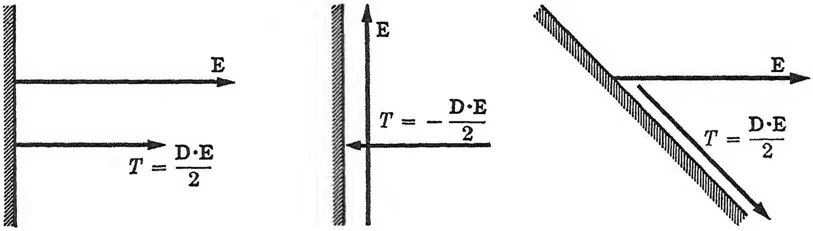

E
$T = \frac{\mathbf{D}\cdot\mathbf{E}}{2}$
E
$T = -\frac{\mathbf{D}\cdot\mathbf{E}}{2}$
E
$T = \frac{\mathbf{D}\cdot\mathbf{E}}{2}$

FIG. 6–3 Stresses for fields perpendicular, parallel, and at an angle of 45° to a surface.

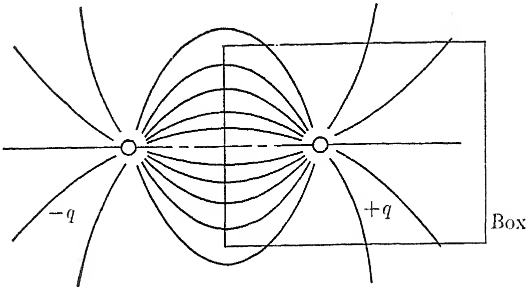


FIG. 6–4 Lines of force for equal and opposite charges.

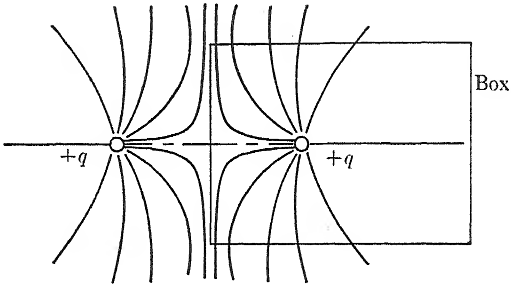


FIG. 6–5 Lines of force for equal charges of the same sign.

If we consider, for example, two charges of equal magnitude but of opposite sign, then the lines of force are distributed as in Fig. 6–4. If the stress tensor is integrated over the surfaces of a box, one of whose faces is the plane of symmetry between the two charges and the other faces are at infinity, the resulting expression is in agreement with the Coulomb attraction of Eq. (1–24). If we consider the two equal like charges, as in Fig. 6–5, and the same box as before, then the lines of force are parallel at the plane of symmetry, resulting in a repulsion whose magnitude can similarly be shown to be also in accord with the Coulomb repulsion.

6–7 The behavior of dielectric liquids in an electrostatic field. Let us consider the behavior of an uncharged dielectric liquid resulting from the volume force produced by an electrostatic field. If p is the mechanical pressure in the liquid when in equilibrium with the electric volume force $\mathbf{F}_v$, then the mechanical volume force $-\nabla p$ which is set up as a result of the pressure gradient is equal and opposite to $\mathbf{F}_v$. In other words, the equilibrium condition is that $\mathbf{F}_v + (-\nabla p) = 0$, so that $\mathbf{F}_v = \nabla p$. Then by Eq. (6–40) the pressure gradient at any point within the liquid is given by

$$\nabla p = \mathbf{F}_v = -\frac{\epsilon_0 E^2}{2}\nabla\kappa + \frac{\epsilon_0}{2}\nabla\left(E^2 g\frac{d\kappa}{dg}\right). \tag{6-58}$$

This equation can be written as

$$\nabla p = \frac{\epsilon_0 g}{2}\nabla\left(E^2\frac{d\kappa}{dg}\right). \tag{6-59}$$

On integrating this, assuming that there is a definite equation of state for the liquid, we obtain

$$\int_{p_1}^{p_2}\frac{dp}{g} = \frac{\epsilon_0}{2}\left\{\left[E^2\frac{d\kappa}{dg}\right]_2 - \left[E^2\frac{d\kappa}{dg}\right]_1\right\}. \tag{6-60}$$

This equation denotes the important fact that the pressure within the dielectric liquid is a unique function of the electric field at a given point, the function depending on the electrical and mechanical equation of state of the liquid. Equation (6-60) also indicates that the net pressure difference (resulting from electrical forces) between two points outside the region of the electric field in a dielectric liquid will vanish. A situation that involves boundaries will be analyzed later.

If the liquid is incompressible, Eq. (6-60) reduces to

$$p_2 - p_1 = \frac{g\epsilon_0}{2}\left[E^2\frac{d\kappa}{dg}\right]_1^2, \tag{6-61}$$

from which the magnitude of the pressure difference can be estimated numerically in terms of the Clausius-Mossotti relation or a similar equation of state. If the Clausius-Mossotti relation [Eq. (2-39)] is valid, Eq. (6-61) for an incompressible fluid becomes

$$p_2 - p_1 = \frac{N_0\epsilon_0 g\alpha}{18M}\left[E^2(\kappa+2)^2\right]_1^2 = \left[\frac{\epsilon_0 E^2}{2}\,\frac{(\kappa-1)(\kappa+2)}{3}\right]_1^2. \tag{6-62}$$

As an example of the stresses that act across the boundary between two dielectrics, let us take the case of a boundary between a dielectric of specific inductive capacity κ and a vacuum. We assume that the transition from dielectric to vacuum takes place in a continuous manner, as indicated by Fig. 6-6. For simplicity, let the problem be a two-dimensional one involving a pair of condenser plates that dip into a

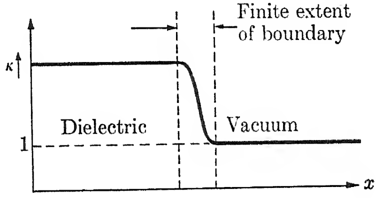


Fig. 6-6 Assumed transition of κ from a dielectric to vacuum.

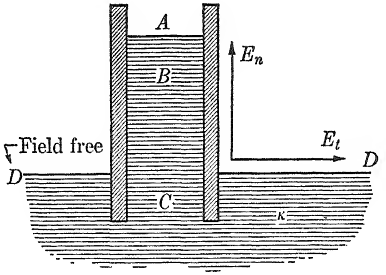


FIG. 6–7 Two condenser plates dipping into a liquid dielectric. The field free point D may be anywhere in the dielectric, well away from the condenser. It is indicated as a surface point to facilitate comparison with the hydrostatic pressure.

dielectric liquid, such as those shown in Fig. 6–7. If the net pressure difference from A to D is all that is desired, it suffices to integrate the one term of Eq. (6–40) that is proportional to the gradient of the dielectric constant. The resultant electrical pressure difference, which has to be balanced by hydrostatic effects, is

$$p_A - p_D = \frac{\epsilon_0}{2}\int_A^D E^2\nabla\kappa \cdot d\mathbf{x}, \tag{6–63}$$

or

$$p_A - p_D = \frac{\epsilon_0}{2}\int_A^D (E_t^2 + E_n^2)\frac{d\kappa}{dx}\,dx. \tag{6–64}$$

If the boundary conditions, Eqs. (2–15) and (2–18), on the normal and tangential components of the electric field are introduced, Eq. (6–64) becomes

$$\begin{aligned} p_A - p_D &= \frac{\epsilon_0}{2}\left[E_{tB}^2(\kappa - 1) + \kappa^2 E_{nB}^2 \int_A^B \frac{d\kappa}{\kappa^2}\right] \\ &= \frac{\epsilon_0(\kappa - 1)}{2}\,[E_t^2 + \kappa E_n^2]_B. \end{aligned} \tag{6–65}$$

Note that the field quantities in Eq. (6–65) refer to the fields inside the liquid. This formula can be used directly to find the rise of the liquid between the condenser plates. It does not describe the detailed pressure behavior of the liquid from A to D, however, since it does not include the electrostriction term. In fact, the pressure change from A to B is actually of opposite sign from the pressure change from A to D. The detailed behavior of the pressure is shown in Fig. 6–8. As the field decreases from B to D the pressure decreases below the outside value at A by an amount which is larger than the pressure rise at the surface A–B. The net differ-

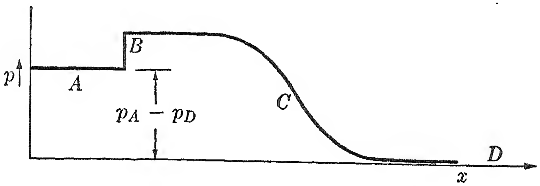


FIG. 6-8 Detailed variation in pressure from A to D. x is simply a path of integration from A to the field free point D.

ence computed in Eq. (6-65) gives only the difference in pressure between A and D. The pressure that forces the liquid up is actually exerted at the region C where the field is inhomogeneous, and not at the surface of the liquid. The physical reason for this is that the energy of dipoles in an electric field is lower than their energy in field-free space, and therefore the dipoles in the liquid are drawn into the regions of higher field in order to satisfy the condition that the potential energy be a minimum. This action on the dipoles takes place in the region C where the field begins to build up. The electrostriction drop at A–B partially counterbalances this minimum energy effect, resulting in the net pressure difference given by Eq. (6-65). The example shows that considerable care is necessary in applying the force equation in dielectrics.

Another example seems extremely simple, but actually leads to an apparent paradox. If a set of charged conductors are so arranged that they may be immersed in a dielectric liquid, and if the true charges on these conductors are kept constant as a liquid is introduced between them, then the free energy of the system as given by Eq. (6-14) will drop by a factor $1/\kappa$, since $\mathbf{D}$ remains constant but $\mathbf{E}$ is reduced. If, on the other hand, the voltages were maintained at their initial values as the liquid was introduced, then the free energy would be increased by a factor κ, since in this case $\mathbf{E}$ remains constant while $\mathbf{D}$ increases. Of course, these arguments pertain only if all of the space between the conductors and outside of them is filled with a dielectric liquid. At least all space where there are electric fields must be so filled, because otherwise we cannot assume that the distribution of $\mathbf{E}$ and $\mathbf{D}$ remains constant as the dielectric material is introduced between the plates.

This means that if a system maintained at constant charge is totally surrounded by a dielectric liquid all mechanical forces will drop in the ratio $1/\kappa$. A factor $1/\kappa$ is frequently included in the expression for Coulomb's law to indicate this decrease in force. The physical significance of this reduction of force, which is required by energy considerations, is often somewhat mysterious. It is difficult to see on the basis of a field theory why the interaction between two charges should be dependent upon the nature or condition of the intervening material, and therefore the inclusion of an extra factor $1/\kappa$ in Coulomb's law lacks a physical explanation.

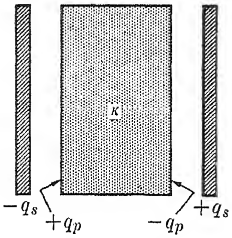


FIG. 6–9 Parallel plate condenser with a slab of solid dielectric.

To investigate this, let us consider a pair of parallel plate conductors between which is inserted a slab of *solid* dielectric, as shown in Fig. 6–9. Let $\pm q_s$ be the surface charge per unit area on each condenser plate and $\pm q_p$ be the polarization charge on the outer surface of the intervening dielectric. The two layers of polarization charge will produce equal and opposite fields on each plate and their effects will therefore cancel each other. From the point of view of electrical interaction alone it is not obvious why any change in force at all is obtained when the dielectric layer is introduced, since the only direct interaction between the charges q_s, which are assumed to remain constant, is unaffected by the introduction of the dielectric slab. That is, the force per unit area remains

$$F_s = \frac{q_s^2}{2\epsilon_0} \tag{6–66}$$

as long as the dielectric does not touch the plates of the condenser. Therefore the decrease in force to

$$F_s = \frac{q_s^2}{2\kappa\epsilon_0}, \tag{6–67}$$

which is experienced when the experiment is performed with a *liquid* that wets the plates and also completely surrounds them, cannot be explained by electrical forces alone.

This apparent paradox can be explained by taking into account the difference in pressure in the liquid between the field-filled space and the field-free space outside the condenser plates. By Eq. (6–65) this difference in pressure is given by

$$p_B - p_A = \frac{\epsilon_0^{-1}}{2}\left(1 - \frac{1}{\kappa}\right) D_n^2 = \frac{q_s^2}{2\epsilon_0}\left(1 - \frac{1}{\kappa}\right). \tag{6–68}$$

The sum of the force resulting from this pressure and the pure electrical force given by Eq. (6–66) gives the total force of Eq. (6–67) which was derived from energy considerations. Thus the decrease in force that is experienced between two charges when they are immersed in a dielectric liquid can be understood only by considering the effect of the pressure of the liquid on the charges themselves. In accordance with the philosophy of the action-at-a-distance theory, no change in the purely electrical interaction between the charges takes place.

SUGGESTED REFERENCES

M. Abraham and R. Becker, *The Classical Theory of Electricity and Magnetism.* Much of our treatment parallels that found in Chapter V of this work.

J. A. Stratton, *Electromagnetic Theory.* The first part of Chapter II is on electrostatic stress and energy.

J. H. Jeans, *Electricity and Magnetism.* Chapter VI, on "the state of the medium in the electrostatic field," begins with a discussion of field *versus* action-at-a-distance, but the detailed treatment is rather brief.

M. Mason and W. Weaver, *The Electromagnetic Field,* especially sections 2 and 37.

EXERCISES

1. Consider a set of conductors at given potentials and carrying a total charge Q. Show that if the charge distribution is perturbed by an infinitesimal amount from that demanded by the field equations, the total energy is increased.

2. Two charges of q at distance d repel each other with a force given by Coulomb's law. Choose a suitable surface surrounding one of the charges and calculate $\mathbf{F}$ by integration of the Maxwell stress tensor.

3. A conducting spherical shell of radius a is placed in a uniform field $\mathbf{E}$. Show that the force tending to separate two halves of the sphere across a diametral plane perpendicular to $\mathbf{E}$ is given by

$$F = \tfrac{9}{4}\pi\epsilon_0 a^2 E^2.$$

4. Consider a long flat conducting strip of width $2a$ with its axis perpendicular to a uniform field $\mathbf{E}$ and its plane inclined at an angle θ to the field. What is the torque per unit length acting on the strip?

5. A dipole of moment $\mathbf{p}$ is placed in a uniform field $\mathbf{E}$. Show by direct integration of the moment of the Maxwell tensor over a sphere centered at the dipole that the torque is equal to the conventional result $\mathbf{p} \times \mathbf{E}$.

6. Find the electrostatic energy of a charge q distributed uniformly over the surface of a sphere of radius r_0. What is the energy if the same charge is distributed uniformly throughout the sphere? (In both cases the energy is of the order $q^2/4\pi\epsilon_0 r_0$.)

7. Find the interaction energy of two interpenetrating spheres of uniform charge density ρ_1 and ρ_2. Take spheres of equal radius, and express the answer in terms of the amount of overlapping or interpenetration.

8. (a) The electric field intensity at a point immediately outside a charged conductor is σ/ϵ_0, where σ is the charge per unit area. Show that half this field arises from charges in the immediate neighborhood and half from distant charges, and hence that the normal stress on the conductor is $\sigma^2/2\epsilon_0$. (b) A piece of gold leaf of area 10 cm^2, weighing one thousandth of a gram, rests on a flat horizontal conducting surface. What electric field intensity must be established at the surface in order to lift the gold leaf?

9. (a) A conducting spherical shell of mass M and radius a floats on oil of specific inductive capacity κ. If a fraction $f_0 < \frac{1}{2}$ of the sphere is immersed while the sphere is uncharged, find the charge q on the sphere which will cause it to become half immersed. In arriving at your answer, show that the dependence of κ on density need not be taken into account. (b) Given all the necessary parameters, what is the pressure *in the liquid* in the immediate neighborhood of the bottom point of the sphere when it is half immersed? What is the force per unit area exerted by the liquid on the sphere at that point?

10. Analyze the fields, pressures, and tensions involved in the situation illustrated: a charged parallel-plate condenser has one plate immersed in a liquid dielectric in such a way that the surface of the dielectric is oblique to the plates (Fig. 6–10).

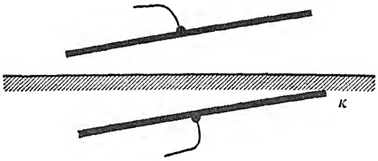


FIG. 6–10

CHAPTER 7

CURRENTS AND THEIR INTERACTION

In the discussion of electrostatics only stationary configurations of charge were under consideration. We shall now take into account the flow of charge, or current. According to Oersted's discovery, every current is accompanied by a magnetic field, and it is thus impossible to treat currents fully by means of electric fields alone. Before investigating magnetic interactions, however, we shall consider currents that vary slowly with the time, and which can be assumed to depend entirely on the electric fields present. In the absence of numerical estimates of the relative magnitude of magnetic interactions and the interaction of currents with the lattice structure of the resistive medium that gives rise to electrical resistance, it is not obvious that a situation ever exists in which currents depend only on electric fields. But it turns out that the magnetic effects can be neglected when the fields vary at low frequencies, provided that the dimensions of the conductor are small compared with the so-called "skin depth" of the currents in the particular conductor. We shall also neglect an additional effect, known as the Hall effect, which is present even at zero frequency, and gives rise to redistribution of the equipotential surfaces in a current-carrying conductor. In all but very special substances, however, this effect is extremely small.

7–1 Ohm's law. The conservation of charge in a medium is expressed by the equation of continuity. If $\mathbf{j}$ is the current density within the medium (measured in amperes/meter^2), this equation is

$$\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0. \tag{7–1}$$

The current is called stationary if there is no accumulation of charge at any point. The criterion for stationary flow is obviously

$$\nabla \cdot \mathbf{j} = 0. \tag{7–2}$$

To relate the theory of current to the theory of the electric field, an equation is necessary which relates the current and field at any particular point in the conducting material. In many cases they are simply proportional, so that

$$\mathbf{j} = \sigma \mathbf{E}. \tag{7–3}$$

Here σ is the electrical conductivity of the material, measured in mho/meter in the mks system. Equation (7–3) is equivalent to Ohm's law.

19741

This relation is a phenomenological characteristic, and is by no means universally valid. The range of current densities over which Eq. (7–3) holds is called the linear range of the particular material, and can be very large, as in metals, or very small, as in a semiconductor. Equation (7–3) implies that the conduction is isotropic. In crystals of low symmetry it must be replaced by a tensor equation.

7–2 Electromotive force. Stationary current is impossible in a purely irrotational electric field, since in a stationary current energy is expended at a rate $\mathbf{j} \cdot \mathbf{E}$ per unit volume, and this energy cannot be provided by an irrotational field. Stationary currents are possible only if there are present sources of electric field known as electromotive forces, which produce nonirrotational fields. If we assume that such electromotive fields exist, and denote them by $\mathbf{E}'$, while $\mathbf{E}$ is the field derivable from a potential, the conduction equation becomes

$$\mathbf{j} = \sigma(\mathbf{E} + \mathbf{E}'). \tag{7–4}$$

We may define the electromotive force $\mathcal{E}$ by

$$\mathcal{E} = \oint (\mathbf{E} + \mathbf{E}') \cdot d\mathbf{l} = \oint \mathbf{E}' \cdot d\mathbf{l} = \oint \frac{\mathbf{j} \cdot d\mathbf{l}}{\sigma}. \tag{7–5}$$

The conservative part of the field, $\mathbf{E}$, drops out of the closed line integration, which means that the current is entirely due to the nonconservative forces, although it is influenced by the conductivity and the geometry.

In case the current density is nearly constant over major portions of the path of integration, as frequently happens, Eq. (7–5) can be written as

$$\mathcal{E} = J \oint \frac{dl}{\sigma S} = JR. \tag{7–6}$$

Here $J = |\mathbf{j}|S$ is the total current, a constant for the circuit (amperes), S is the cross-sectional area of the conductor where the current density is $\mathbf{j}$, and R is the resistance of the conductor (ohms); $\mathcal{E}$ is measured in volts. Equation (7–6) is the form in which Ohm's law is usually stated.

Note that in a case where there is no current, we obtain, by integrating Eq. (7–4) along a line between two points 1 and 2 which traverses all of the region in which there is a nonconservative field,

$$-\int_1^2 \mathbf{E} \cdot d\mathbf{l} = \int_1^2 \mathbf{E}' \cdot d\mathbf{l} = \oint E' \cdot d\mathbf{l} = \mathcal{E}. \tag{7–7}$$

This indicates that the open circuit electrostatic voltage between two points is equal to the total electromotive force in the circuit. It follows that within a region where there are nonconservative fields $\mathbf{E}' = -\mathbf{E}$ in the absence of a current flow. Thus, for example, within a given boundary the nonconservative fields (due, e.g., to the chemical potentials) are exactly equal to the electrostatic field which is set up by the charges on the boundaries, if there is no current.

7-3 The solution of stationary current problems. Formally, the current distribution and the field distribution are entirely defined by the nonconservative field and by the conductivity of the medium. By making use of Eqs. (1-28), (7-2), and (7-4), we can write the expressions for **E** and **j** in terms of **E**′:

$$\begin{aligned} \nabla \cdot \mathbf{j} &= 0, \\ \nabla \times \left(\frac{\mathbf{j}}{\sigma}\right) &= \nabla \times \mathbf{E}', \\ \nabla \cdot (\sigma \mathbf{E}) &= -\nabla \cdot (\sigma \mathbf{E}'), \\ \nabla \times \mathbf{E} &= 0. \end{aligned} \tag{7-8}$$

In the region where there are no nonconservative fields, **E** is derivable from a potential, and hence in the case of stationary flow the potential still obeys Laplace's equation:

$$\mathbf{E} = -\nabla\phi,$$

$$\nabla \cdot (\sigma \nabla \phi) = 0,$$

or $\nabla^2\phi = 0$ if σ is constant. The boundary conditions are different from those of electrostatics, however, since now the conductivities rather than the dielectric constants define the flux relation across a boundary. From Eqs. (7-2) and (7-3) we have, in the absence of nonconservative fields,

$$\nabla \cdot (\sigma \mathbf{E}) = 0. \tag{7-9}$$

Just as in the analogous case for dielectrics in Chapter 2, the divergence equation yields the boundary condition on the normal component of the field at the surface between two media designated by the subscripts 1 and 2:

$$\begin{aligned} \mathbf{n} \cdot (\sigma_2 \mathbf{E}_2 - \sigma_1 \mathbf{E}_1) &= 0, \\ \mathbf{n} \cdot (\sigma_2 \nabla\phi_2 - \sigma_1 \nabla\phi_1) &= 0. \end{aligned} \tag{7-10}$$

The curl equation leads to the condition on the tangential component:

$$\begin{aligned} \mathbf{n} \times (\mathbf{E}_2 - \mathbf{E}_1) &= 0, \\ \mathbf{n} \times (\nabla\phi_2 - \nabla\phi_1) &= 0. \end{aligned} \tag{7-11}$$

It follows that the solution of stationary current distribution problems is mathematically identical to the solution of electrostatic potential distribution problems that have the same geometry, and thus all of the methods developed in Chapters 3, 4, and 5 are applicable to these problems. The only difference between the static current problems and the electrostatic problems is that the conductivity in a given region may be zero, while the specific inductive capacity is in general not less than unity. This means that the type of boundary value problems which arise in stationary current flow may, under certain conditions, be quite different from any that can exist in electrostatic cases. As an example: if we consider that the region between a set of parallel condenser plates is filled with a medium of conductivity σ, then the current in the stationary current range will be exactly uniform over the entire area within the conducting medium, while in the analogous electrostatic case the field distribution will be only approximately uniform and will be disturbed by the fringing field at the edges of the plates.

In general, if electrostatic methods permit the calculation of the capacity between two electrodes, then one can conclude immediately what the resistance will be between those electrodes if all of the space in which they are located is filled with a homogeneous medium. The capacity between two electrodes 1 and 2 is given by

$$C = \frac{\kappa\epsilon_0 \int \mathbf{E} \cdot d\mathbf{S}}{\int_1^2 \mathbf{E} \cdot d\mathbf{l}}. \tag{7-12}$$

The numerator is the charge on each electrode by Gauss's flux theorem, and the denominator is the potential difference between the electrodes. The resistance between the two electrodes is

$$R = \frac{\int_1^2 \mathbf{E} \cdot d\mathbf{l}}{\sigma \int \mathbf{E} \cdot d\mathbf{S}}, \tag{7-13}$$

where the denominator gives the net flux of current and the numerator is the potential difference. On comparing Eqs. (7-12) and (7-13), we see that

$$\frac{1}{R\sigma} = \frac{C}{\kappa\epsilon_0} \tag{7-14}$$

or

$$RC = \frac{\kappa\epsilon_0}{\sigma}. \tag{7-15}$$

We note that the product of the resistance and the capacity is a constant that depends only on the conductivity of the conductor and on the specific inductive capacity of the material between the condenser plates, and not on the geometry. However, it is not always possible to find an electrostatic problem which will be fully analogous to the corresponding stationary current problem, since the difference in the range of dielectric constants and conductivities mentioned above causes different field patterns in the two cases. The formula for the capacity of a parallel plate condenser, when edge effects are neglected, is

$$C = \frac{\kappa\epsilon_0 S}{l}, \tag{7–16}$$

where S is the area of the plates and l is the distance between them. With this substitution for C, Eq. (7–15) gives

$$R = \frac{l}{\sigma S} \tag{7–17}$$

for the resistance of a wire of known length l and cross-sectional area S. The formula for the resistance of a wire is, of course, applicable for large values of l, since the zero conductivity of the surrounding medium prevents fringing of the current flow lines, in strong contrast to the shape of the field lines that would arise from condenser plates that coincided with the ends of a wire of ordinary length.

7–4 Time of relaxation in a homogeneous medium. The equation of continuity (7–1) may be combined with the conductivity equation $\mathbf{j} = \sigma\mathbf{E}$ and the source equation $\nabla \cdot \mathbf{E} = \rho/\kappa\epsilon_0$ to give a differential equation for ρ:

$$-\frac{\partial\rho}{\partial t} = \frac{\sigma\rho}{\kappa\epsilon_0}. \tag{7–18}$$

This equation can be integrated with respect to the time, in case we are dealing with a homogeneous dielectric or a homogeneous conductor. The integration gives

$$\rho = \rho_0 e^{-t/\tau}, \tag{7–19}$$

where the characteristic time τ is

$$\tau = \kappa\epsilon_0/\sigma. \tag{7–20}$$

The characteristic time τ is usually known as the relaxation time of the medium. Note that it is applicable only to the case of a homogeneous medium, for if the medium were not homogeneous the spatial dependence of the conductivity and the dielectric constant would have to be taken into account in the integration of Eq. (7–18). The relaxation time is a characteristic time for a medium in that it gives an indication of the time in which essentially stationary conditions will be reached after the initiation of a particular flow of charge. The criterion that must be used to determine whether the stationary current equations will be applicable in a particular case is whether the time of observation following the inception of such currents exceeds the relaxation time τ by a sufficiently large amount.

7–5 The magnetic interaction of steady currents. The magnetic interaction of currents is best described in terms of an experimentally established interaction in vacuum that is analogous to the electrostatic Coulomb's law. The mathematical generalization of the results of Ampere's experiments, which gave the force between two current-carrying elements, as shown in Fig. 7–1, is

$$\mathbf{F}_2 = \frac{\mu_0}{4\pi} J_1 J_2 \oint_1 \oint_2 \frac{d\mathbf{l}_2 \times (d\mathbf{l}_1 \times \mathbf{r}_{12})}{r_{12}^3}. \tag{7–21}$$

Here μ_0 is a constant characteristic of the system of units, and is equal to $4\pi \times 10^{-7}$ henry/meter in the mks system; $\mathbf{F}_2$ is the force on the circuit that carries the current J_2 and whose line element is $d\mathbf{l}_2$. Due to the geometry that is involved in expressing the relative directions of $\mathbf{F}_2$, $d\mathbf{l}_1$, $d\mathbf{l}_2$, and $\mathbf{r}_{12}$, this force equation appears to be more complicated than the Coulomb force equation, Eq. (1–24). Also it appears, superficially, to violate Newton's third law of the equality of action and reaction. The integrand of Eq. (7–21) is, in fact, asymmetrical as it stands; when the integral is carried out over two closed circuits, however, the resulting force is symmetrical in terms of the geometry of the two interacting current loops.

The symmetry of the force between the two circuits can be shown as follows. If the integrand is expanded by the vector product rule,

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{C}, \tag{7–22}$$

$\mathbf{F}_2$ becomes

$$\mathbf{F}_2 = \frac{\mu_0}{4\pi} J_1 J_2 \oint_1 \oint_2 \left\{ \frac{(d\mathbf{l}_2 \cdot \mathbf{r}_{12})\, d\mathbf{l}_1}{r_{12}^3} - \frac{(d\mathbf{l}_1 \cdot d\mathbf{l}_2)\mathbf{r}_{12}}{r_{12}^3} \right\}. \tag{7–23}$$

In the first term the integrand with respect to $d\mathbf{l}_2$ is an exact differential, since it is the line integral of a gradient, $\oint \nabla(1/r_{12}) \cdot d\mathbf{l}_2$. Therefore this

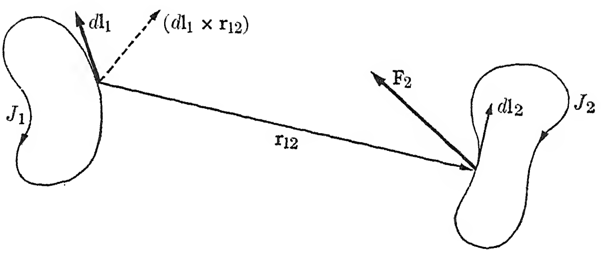


FIG. 7–1 Illustrating Ampere's law.

term vanishes as the integration is carried out over the closed loop. The other integral,

$$\mathbf{F}_2 = -\frac{\mu_0}{4\pi}\left[\oint_1\oint_2 \frac{(d\mathbf{l}_1 \cdot d\mathbf{l}_2)\mathbf{r}_{12}}{r_{12}^3}\right] J_1 J_2, \tag{7-24}$$

is symmetric in terms of loops 1 and 2.

We have here shown that the basic law of interaction, Eq. (7–21), is, in fact, symmetrical in the current elements of a circuit and is therefore not in violation of Newton's third law. The question is often raised as to how contradiction with the third law can be avoided if each current element is considered, not as part of a closed loop, but as a charge moving with velocity **u**. In that case the first term in Eq. (7–23) would not vanish, and therefore Newton's third law is not satisfied. This situation, however, represents nonstationary conditions of electromagnetic field, and one cannot exclude the possibility that momentum will be carried away by the change in electromagnetic field. In fact, we shall show later that electromagnetic momentum has to be associated with the electromagnetic field, and that therefore in a nonstationary problem involving electromagnetic interactions action need not be directly equal to reaction between the two interacting sources. A complete discussion of this point must be postponed until nonstationary problems can be treated consistently.

7–6 The magnetic induction field. The reason for taking Eq. (7–21) as the starting point for the discussion of magnetic interactions is that Eq. (7–21) is in such a form that the interaction expression can be separated into a field produced by loop 1 and a force exerted by this field on loop 2. Equation (7–24), on the other hand, which implicitly contains the cosine of the angle between the circuit elements, does not permit such a separation and therefore does not lead directly to a vector field formulation of magnetic interactions. The separation of Eq. (7–21) into a field and a field force can be carried out by putting

$$\mathbf{F}_2 = J_2 \oint d\mathbf{l}_2 \times \mathbf{B}_2, \tag{7-25}$$

where $\mathbf{B}_2$ is the magnetic field of induction produced by circuit 1 at the position of circuit 2, and is

$$\begin{aligned}\mathbf{B}_2 &= \frac{\mu_0}{4\pi} J_1 \oint_1 \frac{d\mathbf{l}_1 \times \mathbf{r}_{12}}{r^3} \\ &= -\frac{\mu_0}{4\pi} J_1 \oint d\mathbf{l}_1 \times \nabla_2\left(\frac{1}{r_{12}}\right).\end{aligned} \tag{7-26}$$

(The mks unit of **B** is the weber/meter^2.)

B is analogous to **E** in the electrostatic theory in that it determines the force that acts on a circuit element. Eq. (7–26) is a generalization of the law of Biot and Savart. It should be noted that thus far we have no differential form of this law. The Biot and Savart law when expressed in terms of volume currents becomes

$$\mathbf{F} = \int (\mathbf{j} \times \mathbf{B})\, dv, \tag{7–27}$$

$$\mathbf{B} = \frac{\mu_0}{4\pi} \int \frac{\mathbf{j} \times \mathbf{r}}{r^3}\, dv'. \tag{7–28}$$

Note that **r** in Eq. (7–28) is directed from the point of integration or source point (where **j** is located) toward the field point where **B** is being determined.

7–7 The magnetic scalar potential. Let us inquire under what conditions the magnetic induction field **B** can be derived from a scalar potential by the relation

$$\mathbf{B} = -\mu_0 \nabla \phi_m. \tag{7–29}$$

Consider a closed loop carrying a current J, as in Fig. 7–2, together with a field point $(x_\alpha)_1$. If the field **B** were derivable from a scalar magnetic potential ϕ_m, and if the point of observation were moved through a distance $d\mathbf{x}$ from $(x_\alpha)_1$ to $(x_\alpha)_2$, then the increment in the scalar magnetic potential ϕ_m would be given by

$$d\phi_m = -\frac{d\mathbf{x} \cdot \mathbf{B}}{\mu_0}. \tag{7–30}$$

The Biot and Savart expression, Eq. (7–26), may be used for **B**, so that

$$d\phi_m = -\frac{1}{4\pi} J \oint \frac{d\mathbf{x} \cdot (d\mathbf{l} \times \mathbf{r})}{r^3} = -\frac{1}{4\pi} J \oint \frac{\mathbf{r} \cdot (d\mathbf{x} \times d\mathbf{l})}{r^3}. \tag{7–31}$$

(The mixed vector-scalar product permits cyclic permutation.) Equation

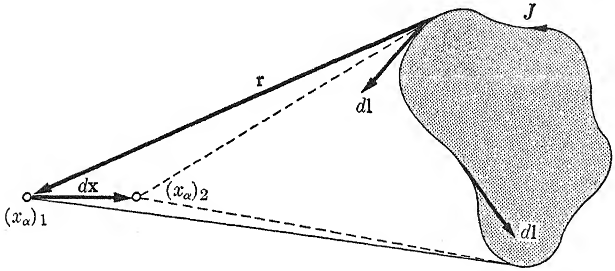


FIG. 7–2 Derivation of magnetic scalar potential.

(7–31) is, of course, equal to the change in the scalar magnetic potential which is obtained if the point of observation were held stationary and the loop were moved by an amount $-d\mathbf{x}$, as was discussed in the derivation of Eq. (1–59) for the potential of a dipole sheet. By means of the relation given in Eq. (1–58), the change in the scalar magnetic potential can be written as

$$d\phi_m = \frac{J}{4\pi}\,d\Omega, \tag{7–32}$$

where $d\Omega$ is the change in the solid angle subtended by the loop at the point of observation brought about by an infinitesimal displacement $-d\mathbf{x}$ of all points of the loop.

This scalar potential has the same mathematical properties as the solution of the electrostatic potential of the surface dipole layer discussed earlier, as seen from the similarity of Eq. (1–54) for the static potential of a dipole layer and Eq. (7–32) for the magnetic induction of a current loop. In addition, this means that the scalar potential of a current loop is multiple valued in the sense that it appears to undergo a discontinuity of magnitude J when a surface bounded by the loop is crossed. In the case of the electric dipole sheet, this surface has a physical significance. In the magnetic case, however, the surface can be chosen in any arbitrary fashion. Since the choice of the surface is arbitrary, the magnetic field derived from such a potential outside the current-carrying region is unambiguous. However, line integrals of the magnetic field of a current loop will be correct only if the path of integration does not pass through the arbitrary surface. The line integral of the magnetic field of induction $\mathbf{B}$ around a closed path threading the current J is exactly equal to the magnitude of the discontinuity in the magnetic scalar potential ϕ_m across the arbitrary reference surface, and hence

$$\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0 J_{\text{total}}. \tag{7–33}$$

For a graphical representation of the magnetic scalar potential of a current loop, see Fig. 1–7. The dipole layer of Fig. 1–7 corresponds to the arbitrarily located surface of discontinuity of the current loop.

Equation (7–33) is the integral representation of the differential relation that gives the total circulation of the magnetic field vector in terms of the current that causes the magnetic field. Since Eq. (7–33) is valid for any arbitrary closed path of integration, we can convert it into a differential expression by substituting $\mathbf{B}$ into Stokes' theorem and reducing the size of the surface of integration to a differential, thus securing

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j}. \tag{7–34}$$

We conclude that **B** cannot in general be derived from a single-valued magnetic scalar potential. The concept of the magnetic scalar potential is of practical utility only to derive magnetic fields in the absence of continuous current distributions. The scalar potential ϕ_m cannot be used if line integrals encircling any currents are considered, or if the fields within current-carrying media are desired.

7–8 The magnetic vector potential. From the defining equation for **B**, Eq. (7–28), we may determine the divergence as well as the curl by direct differentiation:

$$\nabla \cdot \mathbf{B} = \nabla \cdot \left(\frac{\mu_0}{4\pi}\right) \int \frac{\mathbf{j} \times \mathbf{r}}{r^3}\, dv' = \nabla \cdot \left\{\left(\frac{\mu_0}{4\pi}\right)\left[-\int \mathbf{j} \times \nabla\left(\frac{1}{r}\right) dv'\right]\right\}$$

$$= \frac{\mu_0}{4\pi} \int \mathbf{j} \cdot \left[\nabla \times \nabla\left(\frac{1}{r}\right)\right] dv' = 0. \tag{7-35}$$

Similarly,

$$\nabla \times \mathbf{B} = -\frac{\mu_0}{4\pi} \int \nabla \times \left[\mathbf{j} \times \nabla\left(\frac{1}{r}\right)\right] dv'$$

$$= -\frac{\mu_0}{4\pi} \int \mathbf{j} \nabla^2 \left(\frac{1}{r}\right) dv' + \frac{\mu_0}{4\pi} \int [\nabla'(\mathbf{j} \cdot \nabla')] \frac{1}{r}\, dv'.$$

Integrating the second term by parts and dropping a surface term, we have

$$\nabla \times \mathbf{B} = -\frac{\mu_0}{4\pi} \int \mathbf{j} \nabla^2 \left(\frac{1}{r}\right) dv' - \frac{\mu_0}{4\pi} \int (\nabla' \cdot \mathbf{j}) \nabla' \left(\frac{1}{r}\right) dv'$$

$$= \mu_0 \int \mathbf{j}\, \delta(r)\, dv' + 0 = \mu_0 \mathbf{j}, \tag{7-36}$$

in agreement with Eq. (7–34).

We have in the above implicitly assumed that currents are fundamentally the only sources of magnetic field, and that the field of such currents is entirely given by the law of Biot and Savart. From the consequence that the divergence of **B** is zero, it follows that **B** may be derived from a vector potential:

$$\mathbf{B} = \nabla \times \mathbf{A}. \tag{7-37}$$

The vector potential **A** may be identified if we rewrite the expression for **B**, remembering that the operator ∇ acts only on the variable of field position and can therefore be taken outside of the integration sign:

$$\mathbf{B} = -\frac{\mu_0}{4\pi}\int \mathbf{j} \times \nabla\left(\frac{1}{r}\right) dv' = \frac{\mu_0}{4\pi}\nabla \times \int \frac{\mathbf{j}}{r}\, dv'. \tag{7–38}$$

The explicit expression for **A** in terms of the current is therefore given by

$$\mathbf{A} = \frac{\mu_0}{4\pi}\int \frac{\mathbf{j}}{r}\, dv', \tag{7–39}$$

except for an arbitrary function with vanishing curl. For convenience, we may here choose this additional function to be zero, although later this choice will need modification. With this assumption, however, Eq. (7–39) is correct, and the corresponding expression for the vector potential of a linear current distribution is

$$\mathbf{A} = \frac{\mu_0}{4\pi} J \oint \frac{d\mathbf{l}}{r}. \tag{7–40}$$

The fields produced by currents can therefore be computed by first determining the vector potential **A**, using Eq. (7–39) or Eq. (7–40), and then obtaining the magnetic field by means of the relation (7–37). We shall postpone treating the problem of ascertaining vector potentials corresponding to specific current distributions until after we have examined the types of currents encountered and their effect on the fields in material media.

7–9 Types of currents. The original direct experimental observations of the magnetic interactions of currents were made on steady linear currents. Modifications must be introduced into the theory if nonstationary currents are to be treated, and before we can derive the magnetic effects that occur within an arbitrary medium.

We shall here discuss currents in material media in a manner similar to that used to treat charges in material media. Currents may be classified in two categories: true currents that may be identified with the motion of true charges, and other currents which are associated with the medium itself. This separation, which is analogous to the separation that was made in the electrostatic theory between the potentials of true charges and the potentials of polarization charges, will lead us to consider two types of magnetic fields, one derived from true currents and the other derived from the combined effects of all the currents whatever may be their origin. It is this latter field, namely, the field of magnetic induction **B**, that can be considered to be the space-time average of the interatomic fields.

Let us classify the types of currents that must be considered:

1. *True currents,* **j**, identical with the physical transportation of true charges.

2. *Polarization currents*, later shown to be $\partial \mathbf{P}/\partial t$, currents that arise from the change of the polarization with time.

3. *Magnetization currents*, $\mathbf{j}_m$, stationary currents that flow within regions that are inaccessible to observation but which might give rise to net boundary or volume currents, due to imperfect orbit cancellation on an atomic scale.

4. *Convective currents.* If a material medium in motion contains charges of various types additional currents will be obtained which arise from convective effects. These convective currents will be derived from the motion of both true and polarization charges contained in the medium. The convective currents will be discussed in a later chapter. For our present discussion we need a more detailed investigation of polarization and magnetization currents.

7–10 Polarization currents. If ρ represents the charge density within a molecule and the molecular coordinates are designated by $\boldsymbol{\xi}$, then the electrical moment $\mathbf{p}$ of a polarized molecule is defined by

$$\mathbf{p} = \int \mathbf{P}\, dv = \int \rho \boldsymbol{\xi}\, dv, \tag{7–41}$$

where $\mathbf{P} = \rho\boldsymbol{\xi}$. If the charge density within the molecule is changing in time,

$$\frac{\partial \mathbf{p}}{\partial t} = \int \frac{\partial \rho}{\partial t} \boldsymbol{\xi}\, dv. \tag{7–42}$$

By means of the equation of continuity, Eq. (7–42) may be written

$$\frac{\partial \mathbf{p}}{\partial t} = -\int \nabla \cdot (\rho \mathbf{u}) \boldsymbol{\xi}\, dv. \tag{7–43}$$

Integrating by parts and dropping the surface term, which is justified if we choose the surface of integration so that it lies outside the region where there are molecular charges, we obtain

$$\frac{\partial \mathbf{p}}{\partial t} = \int \rho \mathbf{u}\, dv, \tag{7–44}$$

or, on a large scale,

$$\frac{\partial \mathbf{P}}{\partial t} = \frac{\int \rho \mathbf{u}\, dv}{\int dv} = \overline{\rho \mathbf{u}}. \tag{7–45}$$

Hence the quantity $\partial \mathbf{P}/\partial t$ does represent the space-time average value of the molecular currents caused by a varying polarization.

7-11 Magnetization currents. We shall describe these currents in terms of the magnetic moment per unit volume, or magnetization $\mathbf{M}$,

$$\mathbf{M} = \tfrac{1}{2}(\boldsymbol{\xi} \times \mathbf{j}_m), \tag{7-46}$$

where $\boldsymbol{\xi}$ is the coordinate of the current density $\mathbf{j}_m$. The magnetic moment of a particular volume is then

$$\mathbf{m} = \int \mathbf{M}\, dv = \tfrac{1}{2}\int (\boldsymbol{\xi} \times \mathbf{j}_m)\, dv. \tag{7-47}$$

Note that this is analogous to the expression for the mechanical angular momentum in terms of the velocity of a mass distribution in space. It is a purely kinematic definition in the sense that it does not involve any mention of an actual interaction, magnetic or otherwise. This corresponds to the definition of the electric moment of a region given in Eq. (7-41), which is also only a kinematic description of a specific alignment of charges, although the net charge of the volume is zero. The magnetic moment definition is a description of a system of currents which need not produce any net flow across a surface large enough to be accessible to macroscopic observation. In the special case of a single "stationary current" loop that encloses a given area S, $\mathbf{m}$ becomes simply the product of the current in the loop and its area, directed normal to the loop in such a way as to agree with the right-hand rule for the current circulation:

$$\mathbf{m} = J\mathbf{S}. \tag{7-48}$$

This is in agreement with the elementary definition of the magnetic moment of a current loop.

The expression for $\mathbf{j}_m$ in terms of $\mathbf{M}$ follows from an expansion of the vector potential, Eq. (7-39), analogous to that of the scalar potential in Chapter 1. If the current distribution is confined to a volume whose dimensions are small in comparison with the distance to the field point, we may set $\mathbf{r} = \mathbf{R} - \boldsymbol{\xi}$, and expand $1/r$ as in Eq. (1-46). Since the expansion is valid for any current so confined in space, we shall omit the subscript:

$$\frac{4\pi}{\mu_0}\mathbf{A} = \int \frac{\mathbf{j}\, dv}{r} = \frac{1}{R}\int \mathbf{j}\, dv - \left[\frac{\partial}{\partial x_\alpha}\left(\frac{1}{r}\right)\right]_R \int \xi_\alpha \mathbf{j}\, dv$$
$$+ \tfrac{1}{2}\left[\frac{\partial}{\partial x_\alpha}\frac{\partial}{\partial x_\beta}\left(\frac{1}{r}\right)\right]_R \int \xi_\alpha \xi_\beta \mathbf{j}\, dv - \cdots. \tag{7-49}$$

For stationary currents, $\int \mathbf{j}\, dv = 0$, and the vector potential will be given, even in the limit of small distributions of current, by the second term in the expansion. Since $\dfrac{\partial}{\partial x_\alpha}(1/r)$ evaluated at R equals $-R_\alpha/R^3$, we are concerned with

$$\frac{4\pi}{\mu_0} A_\beta = \frac{R_\alpha \int \xi_\alpha j_\beta \, dv}{R^3}, \tag{7-50}$$

where the integration is, of course, over ξ. The integrand of Eq. (7-50) can be transformed by adding and subtracting $\xi_\beta j_\alpha/2$ into

$$\xi_\alpha j_\beta = \tfrac{1}{2}(\xi_\alpha j_\beta + \xi_\beta j_\alpha) + \tfrac{1}{2}(\xi_\alpha j_\beta - \xi_\beta j_\alpha).$$

The integral of the first part of this expression vanishes for a steady current, as may be seen by writing

$$\xi_\alpha \xi_\beta (\nabla \cdot \mathbf{j}) = 0 = \nabla \cdot (\xi_\alpha \xi_\beta \mathbf{j}) - \mathbf{j} \cdot \nabla(\xi_\alpha \xi_\beta).$$

If we apply Gauss's theorem to the complete divergence, we obtain

$$0 = \int \mathbf{j} \cdot \nabla(\xi_\alpha \xi_\beta) \, dv = \int (\xi_\alpha j_\beta + \xi_\beta j_\alpha) \, dv.$$

But

$$R_\alpha(\xi_\alpha j_\beta - \xi_\beta j_\alpha) = [(\boldsymbol{\xi} \times \mathbf{j}) \times \mathbf{R}]_\beta,$$

so that

$$\frac{4\pi}{\mu_0} \mathbf{A} = \int \frac{(\boldsymbol{\xi} \times \mathbf{j}) \times \mathbf{R}}{2R^3} \, dv = \frac{\mathbf{m} \times \mathbf{R}}{R^3} = -\mathbf{m} \times \nabla\left(\frac{1}{R}\right). \tag{7-51}$$

Hence the leading term of the multipole expansion yields the vector potential of a magnetic dipole **m**. If we have a density of magnetization **M**, the vector potential becomes

$$\frac{4\pi}{\mu_0} \mathbf{A} = +\int \mathbf{M} \times \nabla'\left(\frac{1}{r}\right) dv = \int \frac{\nabla' \times \mathbf{M}}{r} \, dv - \int \nabla' \times \left(\frac{\mathbf{M}}{r}\right) dv. \tag{7-52}$$

But

$$\int \nabla' \times \frac{\mathbf{M}}{r} \, dv = -\int \frac{\mathbf{M} \times d\mathbf{S}}{r}.$$

If the surface may be taken outside the region where there is current, then the second term of Eq. (7-52) vanishes. From the general form of the vector potential $\nabla \times \mathbf{M}$ can therefore be identified as the current density and, in particular,

$$\mathbf{j}_m = \nabla \times \mathbf{M} \tag{7-53}$$

enables us to express a volume magnetization in terms of an equivalent current.

The equivalence of magnetized materials and currents in their effect on outside points was first noted by Ampere. The concept is necessary for treating the inaccessible currents of atomic origin in a macroscopic theory.

According to Eq. (7–53), the microscopic currents cancel within a region of homogeneous magnetization; $\mathbf{j}_m$ is the net current density produced where there is inhomogeneous magnetization.

Physically, Eq. (7–53) can be understood as follows. Consider the z-component of the magnetic moment of neighboring current loops in a rectangular network in the x,y-plane, as seen in Fig. 7–3. If the magnetization is inhomogeneous, there will not be complete cancellation between the boundaries of the loops and there will be a net current. This net current will bring about whatever net effects are ascribable to the currents. The moment of a single loop in Fig. 7–3 is

$$\mathbf{m} = \mathbf{M}\, dx\, dy\, dz.$$

Then, from Eq. (7–48), the current in rectangle 1 is

$$J_1 = \frac{M_z\, dx\, dy\, dz}{dx\, dy},$$

and by means of Taylor's expansion we may write the current in the neighboring rectangle 2:

$$J_2 = \frac{[M_z + (\partial M_z/\partial x)\, dx]\, dx\, dy\, dz}{dy\, dx}.$$

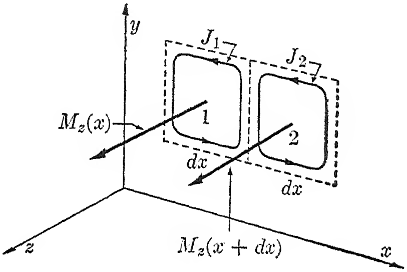


FIG. 7–3 Equivalent Amperian current loops in a magnetized medium, showing cancellation effect on internal boundary.

The difference between J_1 and J_2 results in a net current in the y-direction along the mutual boundary of rectangles 1 and 2:

$$J_y = -\frac{\partial M_z}{\partial x}\, dx\, dz.$$

This form will be recognized as one of the six components of the curl and, in general,

$$\mathbf{j}_m = \nabla \times \mathbf{M},$$

in agreement with Eq. (7–53). In a region of discontinuous magnetization it is easily seen that a surface current equal to the change in the tangential component of the magnetization will result at such a discontinuity. This follows when Eq. (7–53) is applied to a limiting transverse surface bounding such a discontinuity.

7–12 Vacuum displacement current. In a stationary medium the total current is given by the sum of the first three types of current enumerated in Section 7–9:

$$\mathbf{j}_{\text{total}} = \mathbf{j}_{\text{true}} + \frac{\partial \mathbf{P}}{\partial t} + \nabla \times \mathbf{M}. \tag{7–54}$$

To assure the conservation of charge, it is necessary that this total current obey the equation of continuity, $\nabla \cdot \mathbf{j} + (\partial\rho/\partial t) = 0$. The divergence of the total current is

$$\nabla \cdot \mathbf{j}_{\text{total}} = \nabla \cdot \mathbf{j}_{\text{true}} + \nabla \cdot \left(\frac{\partial \mathbf{P}}{\partial t}\right) + \nabla \cdot (\nabla \times \mathbf{M}), \tag{7–55}$$

so that the equation of continuity as applied to the total charge and current densities is

$$\nabla \cdot \mathbf{j}_{\text{true}} + \nabla \cdot \left(\frac{\partial \mathbf{P}}{\partial t}\right) + \nabla \cdot (\nabla \times \mathbf{M}) + \left(\frac{\partial \rho}{\partial t}\right)_{\text{total}} = 0. \tag{7–56}$$

Since

$$\rho_{\text{total}} = \epsilon_0 \nabla \cdot \mathbf{E},$$

$$\left(\frac{\partial \rho}{\partial t}\right)_{\text{total}} = \epsilon_0 \nabla \cdot \frac{\partial \mathbf{E}}{\partial t}, \tag{7–57}$$

and thus

$$\nabla \cdot \mathbf{j}_{\text{true}} + \nabla \cdot \left(\frac{\partial \mathbf{P}}{\partial t}\right) + \nabla \cdot (\nabla \times \mathbf{M}) + \epsilon_0 \nabla \cdot \left(\frac{\partial \mathbf{E}}{\partial t}\right) = 0, \tag{7–58}$$

or

$$\nabla \cdot \left(\mathbf{j}_{\text{true}} + \frac{\partial \mathbf{D}}{\partial t} + \nabla \times \mathbf{M}\right) = 0, \tag{7–59}$$

if we make use of the relation $\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}$.

The divergence of the total current, Eq. (7–55) is not zero, and thus the total current is not solenoidal. However, the quantity

$$\mathbf{c} = \mathbf{j}_{\text{true}} + \frac{\partial \mathbf{D}}{\partial t} + \nabla \times \mathbf{M} = \mathbf{j}_{\text{total}} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}, \tag{7–60}$$

generated by adding the term $\epsilon_0(\partial \mathbf{E}/\partial t)$ to the total current, is a solenoidal current. The need for the addition of this term to produce a solenoidal net current vector was recognized by Maxwell. The "vacuum displacement current" $\epsilon_0(\partial \mathbf{E}/\partial t)$ does not have the significance of a current in the sense of being the motion of charges. We shall see later that the magnetic effects of currents can be formulated only in terms of solenoidal currents, and therefore that the vacuum displacement current term must be introduced in order to be able to apply the formulas which will be developed for the magnetic interaction to cases involving nonstationary currents.

The geometrical significance of the solenoidal current **c** is that at points where there is an accumulation of charge the current is assumed to be continuous across the discontinuity in the form of the rate of change of the field resulting from the accumulation of the charges on the boundaries of the discontinuity. As an example, a battery charging a condenser produces a closed current loop in terms of **c**.

SUGGESTED REFERENCES

M. Abraham and R. Becker, *The Classical Theory of Electricity and Magnetism.* Applications are not emphasized, but the approach to steady currents in Chapter VI is much the same as ours.

G. P. Harnwell, *Principles of Electricity and Electromagnetism.* Stresses linear circuit theory.

W. R. Smythe, *Static and Dynamic Electricity.* Contains many useful and interesting applications.

J. A. Stratton, *Electromagnetic Theory.* Discusses conduction problems along with the electrostatic field.

J. H. Jeans, *Electricity and Magnetism.* The order of material is so different from ours that Jeans may be of limited use in the matter of physical principles, but there are interesting applications.

EXERCISES

1. Calculate the change in resistance of a wire of diameter D (expressed as a change in length of the wire) due to an internal fissure of diameter d. Let $D - d \ll D$, and treat as a two-dimensional problem. (See Fig. 7–4.)

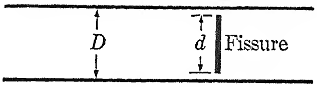


Fig. 7–4

2. The core (radius a) of a coaxial cylindrical cable is surrounded by an insulating sheath of conductivity σ_1, outer radius b, and a second layer of conductivity σ_2 extending to the outer metallic conductor of radius c. Find the resistance per meter of cable between the core and the outer conductor.

3. Current enters an infinite plane conducting sheet at some point p and leaves at infinity. A circular hole, exclusive of p, is cut anywhere in the sheet. Show that the potential difference between any two points on the edge of the hole is twice that between the same two points before the hole was cut.

4. Two small spherical electrodes of radius a are embedded in a semi-infinite medium of conductivity σ, each at distance $d > a$ from the plane face of the medium and at distance b from each other. Find the resistance between the electrodes.

5. A spherical electrode of radius a is surrounded by a concentric spherical electrode of radius b, while the intervening space is filled with a medium whose conductivity is inversely proportional to the distance from the center of the system. If the outer sphere is maintained at potential ϕ_0 and a total current J flows between the electrodes, find the potential at a distance $r > a$ from the center.

6. The surface of a circular disk of radius a is covered with a continuous and uniform spiral winding of N turns of fine wire, starting at the center and continuing to the edge, through which flows a steady current J. Find the magnetic dipole moment.

7. A circular loop, $z = b, x^2 + y^2 = a^2$, carries current J. Show that for $r < R = \sqrt{a^2 + b^2}$ the magnetic scalar potential is proportional to

$$\sum_{n=1}^{\infty} \frac{n+1}{2n+1}\left[P_{n+1}\left(\frac{b}{R}\right) - P_{n-1}\left(\frac{b}{R}\right)\right]\left(\frac{r}{R}\right)^n P_n(\cos\theta).$$

CHAPTER 8

MAGNETIC MATERIALS AND BOUNDARY VALUE PROBLEMS

Thus far only magnetic fields in a vacuum have been treated, although we have analyzed the currents which will have to be taken into account if material media are introduced into magnetic fields. In general the current $\mathbf{j}$ in Eq. (7–34) must be replaced by the total current including the magnetization and polarization currents, as given in Eq. (7–55). We have seen, however, that *in vacuo* the curl of $\mathbf{B}$ is proportional to the total stationary current density, while $\mathbf{j}_{\text{total}}$ is not solenoidal if the polarization changes with the time. The relation $\nabla \times \mathbf{B} = \mu_0 \mathbf{j}$ can be generalized to nonstationary cases in one of two ways: either the current remains solenoidal or the relation used in deriving the magnetic field from the current is modified. The choice between these alternatives made by Maxwell was to retain the relations that derive the magnetic field from the current, Eqs. (7–26), (7–28), or (7–38), but to use the general current $\mathbf{c}$ of Eq. (7–60), which includes the displacement current and which remains solenoidal. This choice has been amply justified by its further consequences.

8–1 Magnetic field intensity. If $\mathbf{c}$ is used as the total current the defining equations for the vector field $\mathbf{B}$ are

$$\nabla \cdot \mathbf{B} = 0, \tag{8–1}$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{c} = \mu_0 \left(\mathbf{j}_{\text{true}} + \nabla \times \mathbf{M} + \frac{\partial \mathbf{D}}{\partial t} \right). \tag{8–2}$$

In the treatment of the polarization of dielectrics in Chapter 2 it was found mathematically convenient to separate the field whose sources were true charges from the total field whose sources were the true charges plus the polarization charges. In a similar manner, it is convenient to separate from the total field that part whose circulation density arises from atomic magnetization currents. Therefore if we write Eq. (8–2) in the form

$$\nabla \times (\mathbf{B} - \mu_0 \mathbf{M}) = \mu_0 \left(\mathbf{j}_{\text{true}} + \frac{\partial \mathbf{D}}{\partial t} \right) \tag{8–3}$$

and define a new field $\mathbf{H}$ by

$$\mathbf{H} = \frac{1}{\mu_0} (\mathbf{B} - \mu_0 \mathbf{M}) = \frac{\mathbf{B}}{\mu_0} - \mathbf{M}, \tag{8–4}$$

this definition reduces to

$$\nabla \times \mathbf{H} = \mathbf{j}_{\text{true}} + \frac{\partial \mathbf{D}}{\partial t}. \tag{8–5}$$

The quantity **H**, the magnetic field intensity, is measured in amp-turns/m in the mks system of units. Equation (8–5) then means that the circulation density of **H** arises from the true current plus the total displacement current, the latter including both the polarization current $\partial \mathbf{P}/\partial t$ and the vacuum displacement current $\epsilon_0\, \partial \mathbf{E}/\partial t$. Under stationary conditions, or quasi-stationary conditions in which the magnetic effect of the displacement current is negligible compared with the magnetic effect of the true current,

$$\nabla \times \mathbf{H} = \mathbf{j}_{\text{true}}, \tag{8–6}$$

and in integral form,

$$\oint \mathbf{H} \cdot d\mathbf{l} = J_{\text{true}}. \tag{8–7}$$

Note that in the sense of the separation between the effect that is produced by the total and the true charges and the total and true currents, respectively, **B** plays a role that corresponds to **E**, while **H** plays a role that corresponds to **D**, as may be seen by comparing the above equations with

$$\int \mathbf{E} \cdot d\mathbf{S} = \frac{q_{\text{total}}}{\epsilon_0},$$

$$\int \mathbf{D} \cdot d\mathbf{S} = q_{\text{true}}$$

of Chapter 2.

8–2 Magnetic sources. The discussion of dielectrics in Chapter 2 was limited to the case of linear media, namely, media in which the polarization was proportional to the applied electric field. In ferromagnetic substances, however, the case of nonlinear behavior is most common, and therefore we must discuss some of the properties of the magnetic field which arise in cases when the magnetization **M** is not a linear and often not even a unique function of the external fields. At first we shall assume **M** to be a given function of the material medium independent of external fields. In the most extreme case, that of a permanent magnet, there will be a magnetic moment **M** per unit volume, even in the absence of any true currents. In this case, as we see from Eq. (8–6), **H** is irrotational and will therefore behave mathematically like an electrostatic field, while **B** remains, of course, solenoidal:

$$\nabla \times \mathbf{H} = 0,$$
$$\nabla \times \mathbf{B} = \mu_0(\nabla \times \mathbf{M}) \neq 0. \tag{8–8}$$

However, the magnetic field will have sources; if we call the magnetic source density ρ_m,

$$\nabla \cdot \mathbf{H} = -\nabla \cdot \mathbf{M} = \rho_m. \tag{8-9}$$

The "magnetostatic field" **H** can therefore be derived from a magnetic source density which is equal to the negative divergence of the magnetization. One unit of this equivalent magnetic charge density is usually known as a unit magnetic pole. In terms of this description, a magnetic pole has no physical reality other than that it makes the mathematical description of the resultant magnetic field of a permanent magnet formally the same as that of the electric field of charges. Since the magnetic field **H** of a permanent magnet is irrotational, it can be derived from a magnetic scalar potential ϕ_m in the same way that **E** may be derived from the electrostatic potential ϕ. If we put

$$\mathbf{H} = -\nabla\phi_m, \tag{8-10}$$

then the resultant scalar potential, in terms of the equivalent volume and surface pole densities, is given by

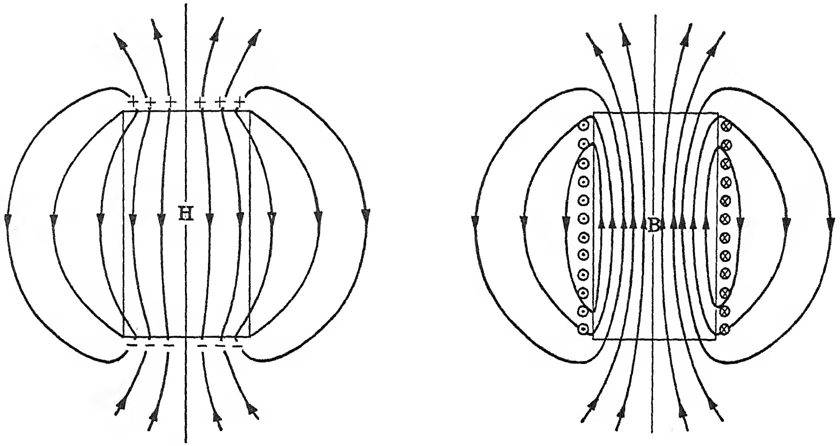
$$\phi_m = \frac{1}{4\pi}\left[\int \frac{\mathbf{M} \cdot d\mathbf{S}}{r} - \int \frac{\nabla \cdot \mathbf{M}}{r}\, dv\right]. \tag{8-11}$$

The field of a permanent magnet with a given magnetization can also be described by the vector potential **A** which is derived from the equivalent surface *currents* and volume *currents* within the magnetized body. Equation (7–52) becomes

$$\mathbf{A} = \frac{\mu_0}{4\pi}\left[\int \frac{\nabla \times \mathbf{M}}{r}\, dv - \int \frac{\mathbf{n} \times \mathbf{M}}{r}\, dS\right], \tag{8-12}$$

where the surface now coincides with the boundary of the magnet. The surface current is equivalent to $-\mathbf{n} \times \mathbf{M}$, where **n** is a unit vector normal to the surface. The magnetic induction field is then derived from Eq. (8–12) by the use of Eq. (7–37).

In the case of a uniformly magnetized medium all internal currents cancel and hence the equivalent surface currents are the only ones present. A cylindrical magnet magnetized in a direction parallel to the axis of the cylinder therefore has a magnetic induction field equivalent to the field of a solenoidal coil, carrying current on the cylindrical face of the magnet, with the current flow lines lying in planes normal to the axis of the cylinder. In this case it is clear that contributions to the scalar potential of Eq. (8–11) come only from the ends of the cylinder. The situation can be described by noting that for a permanent magnet **H** can be thought of as arising from a layer of equivalent pole charges located on the magnet pole



(a)

(b)

Fig. 8–1 A permanent magnet of uniform magnetization: (a) showing equivalent magnetic charges from which **H** is derived, and (b) showing equivalent solenoid with resultant **B**.

faces in the same manner as an electrostatic field would be formed by charges so placed. On the other hand, **B** arises from an equivalent solenoid which can be thought of as being wound on the cylindrical surface of the magnet in the same manner as a vacuum steady current field arises. $\mathbf{B}/\mu_0$ and **H** are identical outside of the region where **M** has a finite value, but they differ by **M** inside the magnet. Note that **B** and **H** in Fig. 8–1 are actually in opposite directions inside the magnet, as must be true if the line integral of **H** is to be zero over any closed path.

Equations (8–11) and (8–12) give scalar and vector potentials in terms of the equivalent pole or current distributions. The potentials can, of course, be described in terms of the integral over the potentials of the individual magnetic moments themselves. We have already seen in Eq. (7–51) that

$$\mathbf{A} = -\frac{\mu_0}{4\pi}\int \mathbf{M} \times \nabla\left(\frac{1}{r}\right) dv'. \tag{8–13}$$

Similarly, the scalar potential can be written in a form analogous to Eq. (1–61):

$$\phi_m = -\frac{1}{4\pi}\int \mathbf{M} \cdot \nabla\left(\frac{1}{r}\right) dv'. \tag{8–14}$$

The equivalence of these expressions with Eqs. (8–11) and (8–12) is evident from an integration by parts. The fields derived from either set of equations must, of course, be the same.

By the use of vector identities it can be shown that the vacuum field due to a magnetic dipole of moment $\mathbf{m} = \int \mathbf{M}\, dv'$ is

$$\mathbf{B} = \nabla \times \mathbf{A} = -\frac{\mu_0}{4\pi}\left\{\nabla \times \left[\mathbf{m} \times \nabla\left(\frac{1}{r}\right)\right]\right\}$$

$$= \frac{\mu_0}{4\pi}\left[(\mathbf{m}\cdot\nabla)\nabla\left(\frac{1}{r}\right) - \mathbf{m}\nabla^2\left(\frac{1}{r}\right)\right]. \tag{8–15}$$

The last term in Eq. (8–15) vanishes except at $r = 0$. From Eq. (8–14),

$$\mathbf{H} = -\nabla\phi_m = \frac{1}{4\pi}\left\{\nabla\left[\mathbf{m}\cdot\nabla\left(\frac{1}{r}\right)\right]\right\} = \frac{(\mathbf{m}\cdot\nabla)\nabla\left(\frac{1}{r}\right)}{4\pi}. \tag{8–16}$$

Thus the two fields differ, except for a factor μ_0, only at $r = 0$.

We have now seen that permanent magnets may be described either in terms of "equivalent currents" or "equivalent poles." Since the entire description of magnetic fields has been based on the premise that they are produced by moving charges* we are led to believe that the interpretation of the field of a permanent magnet in terms of the circulation of atomic currents is a more fundamental one than the concept of magnetic charges, and that therefore **B**, which arises from currents, is a more fundamental field than **H**, which arises from "magnetic charges." The description in terms of **H** is more attractive from a practical point of view, however, since it reduces problems that involve permanent magnets or problems involving magnetized pieces of iron whose magnetization can be determined by other means, to problems in electrostatics.

The question as to whether **B** or **H** is the more fundamental field can be formulated in a different way. Let us pose the problem: "Consider a charge q moving with a velocity $\mathbf{u}$ in a magnetized medium and let us suppose that the force acting on it is of the form $\mathbf{F} = q(\mathbf{u} \times \mathbf{X})$. Should we use **B** for **X**, or should we use **H** or even a combination of the two?" This question was first tested experimentally by Rasetti† by measuring the deflection of cosmic rays in magnetized iron, and has been studied theoretically by Wannier‡ by analyzing in detail the motion of charged particles in magnetized media. The answer is essentially this: If the motion of the charged particle is truly random relative to the magnetized material—that is, if it is not affected by the presence of the magnetized medium, to a first approximation—then the force that is exerted on it corresponds to the use of **B** as the magnetic field in the force equation. If, on the other hand, the particle is moving slowly and its motion is substantially affected by the magnetized medium, then it is effectively prevented from passing through the equivalent atomic current loops, and in this case, since the individual current loops act like impenetrable dipoles, the averaging process favors a deflection that corresponds to the use of **H** in the force equation. Rasetti's experimental results actually indicate that the deflection for very high speed particles corresponds approximately to the use of **B** in the force equation. In order for the results to correspond to the use of **H** it would have been necessary for the deflection to have been in the opposite direction. The answer to the above question can be given precisely in the limit of high velocities: $\mathbf{X}_{\mathrm{u}\to\mathrm{c}} = \mathbf{B}$.

---

* There is no basic objection to the existence of magnetic poles; their fields are simply not considered here since there is no experimental evidence as to their existence. If single magnetic poles did exist all the above equations would have to be supplemented. It has been shown quantum mechanically that if magnetic poles did exist the magnitude of the "elementary" unit pole would have to be related to the inverse of the elementary charge by a constant factor.

† F. Rasetti, *Phys. Rev.* **66,** 1 (1944).

‡ G. H. Wannier, *Phys. Rev.* **72,** 304 (1947).

Further evidence for the basic importance of **B** has been obtained from experiments on the reflection and diffraction of neutrons by magnetic materials performed by Hughes and others. It has been shown that the effective field for neutrons is **B**, not **H**; that is to say, the neutrons and the magnetic domains interact as if both were Amperian currents, not impenetrable dipoles. A discussion of the relevant experiments and their interpretation is to be found in *Pile Neutron Research*, by D. J. Hughes.*

8–3 Permeable media: magnetic susceptibility and boundary conditions.

Thus far we have considered the magnetization **M** as a given function of position, as in a permanent magnet. We now turn to the case of an ideally permeable medium, i.e., a medium that has no magnetic moment in the absence of external true currents, and in which there is a magnetic moment proportional to the field produced by any external true currents. The field equations are then

$$\nabla \cdot \mathbf{B} = 0 \tag{8–1}$$

and

$$\oint \mathbf{H} \cdot d\mathbf{l} = J_{\text{true}}. \tag{8–7}$$

If we assume that $\mathbf{M} = \chi_m \mathbf{H}$, we get a relation that corresponds to Eq. (2–9) in the discussion of electrostatics:

$$\mathbf{B} = \mu_0(\chi_m + 1)\mathbf{H} = \kappa_m \mu_0 \mathbf{H} = \mu \mathbf{H}, \tag{8–17}$$

where χ_m is the magnetic susceptibility, $\kappa_m = \chi_m + 1$ is the relative permeability of the medium, and μ is its absolute permeability.

---

* D. J. Hughes, *Pile Neutron Research*, Addison-Wesley, 1954, particularly Sections 11–4 and 10–6.

By means of a derivation which is completely analogous to the one used in Chapter 2 to derive the boundary conditions for **E** and **D**, the boundary conditions for **B** and **H** may be found. At a boundary between two linear media 1 and 2 the relation between the normal components is

$$\mathbf{n}_1 \cdot (\mathbf{B}_2 - \mathbf{B}_1) = \mathbf{n}_1 \cdot (\mu_2 \mathbf{H}_2 - \mu_1 \mathbf{H}_1) = 0, \tag{8-18}$$

while for the tangential components,

$$\mathbf{n}_1 \times (\mathbf{H}_2 - \mathbf{H}_1) = \mathbf{n}_1 \times (\nabla\phi_{m_1} - \nabla\phi_{m_2}) = \mathbf{n}_1 \times \left(\frac{\mathbf{B}_2}{\mu_2} - \frac{\mathbf{B}_1}{\mu_1}\right) = \mathbf{K}, \tag{8-19}$$

where **K** is the true surface current on the boundary between the two media. Equation (8–18) is analogous to Eq. (2–14) and Eq. (8–19) to Eq. (2–18), but there are differences in each case. The normal component of **B** is strictly continuous across the boundary, while the normal component of **D** is continuous only if there is no surface charge. On the other hand, the tangential component of **E** is strictly continuous, while that of **B** is continuous across the boundary only if there is no surface current. These differences correspond, of course, to the existence of true electric charge and the absence of true magnetic charge.

8–4 Magnetic circuits. It may be noted that Eqs. (8–1), (8–17), and (8–7) are mathematically identical with the equations governing stationary flow in a continuous medium in the presence of a nonconservative electromotive force, which were

$$\nabla \cdot \mathbf{j} = 0, \tag{7-2}$$

$$\mathbf{j} = \sigma \mathbf{E}, \tag{7-3}$$

$$\oint \mathbf{E} \cdot d\mathbf{l} = \mathcal{E}. \tag{7-5}$$

These last three equations led to the expression

$$R = \frac{l}{\sigma S} = \sum_i \frac{l_i}{\sigma S_i} \tag{7-17}$$

for the "resistance" of linear conductors in series. This analogy gives rise to the concept of the magnetic circuit, and the solution of linear magnetic media problems given by the expression for the magnetic flux, Φ_m:

$$\Phi_m = \int \mathbf{B} \cdot d\mathbf{S} = \frac{J}{R_m}, \tag{8-20}$$

where

$$R_m = \sum_i \frac{l_i}{\mu S_i} \tag{8-21}$$

is the magnetic "reluctance" (ampere/weber) of the circuit. This solution is based solely on the correspondence of the differential equations for linear magnetization problems with those for steady current problems, and the solutions themselves will actually correspond only in case the boundary conditions for the magnetic and current problems are identical. This cannot be true in general. In fact, it can never be completely accurate, since the conductivity of free space is zero, while the permeability of free space is unity. This means that the magnetic circuit solution will be valid only if the permeability of the media being considered is large compared with unity, or if the regions of space accessible to the magnetic field which have $\kappa_m \simeq 1$ are small compared with those regions in which the permeability is much larger. The set of solutions (8–20) and (8–21) do, however, form the basis of industrial magnetic machinery design, since they permit an approximate treatment in cases where direct boundary value solutions are impractical.

8–5 Solution of boundary value problems by magnetic scalar potentials. In general, boundary value problems involving magnetic media can be attacked by the use of either the magnetic scalar or the vector potential. Problems involving magnetic media located in external fields, where true currents do not enter the region of interest, are best treated by means of magnetic scalar potentials. In this case the magnetic boundary conditions, Eqs. (8–18) and (8–19), may be expressed in terms of the scalar potentials, and are completely analogous to the electrostatic boundary conditions of Chapter 2 if the relative permeability replaces the specific inductive capacity.

For example, problems involving magnetic shields can be treated by electrostatic boundary value methods. The only additional complication which enters in the magnetic problem is the fact that in practical cases the permeability μ for a particular material is not constant for all values of the external field. This is especially true for high flux densities within the permeable medium. In such cases of high flux density or saturation the method of successive approximations may be used if the permeability is known as a function of the flux density. A solution based on the assumption that μ is constant is first obtained. The resultant flux density is computed, after which the problem is repeated, using the corresponding permeability. This method will give an accurate solution if the field inside the permeable medium turns out to be uniform. The problem of an ellipsoid of magnetic material situated in a uniform magnetic field is such a problem. For this reason, the torque acting on an ellipsoid suspended in a uniform field can be used as a measurement of the permeability of the material of the ellipsoid as a function of the external field. Problems of the behavior of permeable media in high fields where the resultant magnetization is appreciably nonuniform are essentially impossible to treat by purely analytical methods.

8–6 Uniqueness theorem for the vector potential. Problems in which currents are present must be treated by the use of the vector potential unless it is possible to introduce an equivalent dipole sheet in place of the current. We have seen in Section 1–1 that the vector potential is unique if the integration over sources extends to infinity and the sources are confined to a finite region of space. In order to use it with confidence in practical problems we must investigate its uniqueness for a region with definite boundaries. The proof is very similar to the proof used in Chapter 3 for the uniqueness of the scalar potential. We shall need the vector form of Green's theorem. Let $\mathbf{U}$ and $\mathbf{W}$ be arbitrary vector functions, and substitute

$$\mathbf{V} = \mathbf{U} \times (\nabla \times \mathbf{W})$$

into Gauss's divergence theorem, $\int \nabla \cdot \mathbf{V}\, dv = \int \mathbf{V} \cdot d\mathbf{S}$. We obtain

$$\int [(\nabla \times \mathbf{W}) \cdot (\nabla \times \mathbf{U}) - \mathbf{U} \cdot \nabla \times (\nabla \times \mathbf{W})]\, dv = \int \mathbf{U} \times (\nabla \times \mathbf{W}) \cdot d\mathbf{S}, \tag{8–22}$$

which is the required theorem. Let us now consider a region in space that is bounded by the surface S, on which the vector potential is specified, and which has a volume v within which there is no current. In order to accomplish this it may be necessary to choose subsurfaces which will exclude the regions of surface flow. If we now let $\mathbf{U} = \mathbf{W} = \mathbf{A}$ for all points within v, and remember that if $\mathbf{j}_{\text{total}} = 0$ then

$$\nabla \times \mathbf{B} = \nabla \times (\nabla \times \mathbf{A}) = 0,$$

we obtain, on substitution in Eq. (8–22),

$$\int (\nabla \times \mathbf{A})^2\, dv = \int (\mathbf{A} \times \mathbf{B}) \cdot d\mathbf{S} = \int (\mathbf{A} \times \mathbf{B}) \cdot \mathbf{n}\, dS. \tag{8–23}$$

Equation (8–23) may be written in the form

$$\int (\nabla \times \mathbf{A})^2\, dv = \int (\mathbf{B} \times \mathbf{n}) \cdot \mathbf{A}\, dS = \int \mathbf{B}_t \cdot \mathbf{A}\, dS, \tag{8–24}$$

where $\mathbf{B}_t$ is the tangential component of $\mathbf{B}$, parallel to the surface and perpendicular to $\mathbf{S}$. Now let us assume that $\mathbf{A}$ in Eqs. (8–23) and (8–24) represents the difference between alternative solutions corresponding to the same boundary values of either the tangential component of the magnetic induction field $\mathbf{B}$ or the tangential component of vector potential $\mathbf{A}$. The right side of Eq. (8–24) then vanishes, since it is evaluated on the boundary where the alternate solutions are equal. On the other hand, the left side of Eq. (8–24) is positive definite and hence the integrand must vanish. Therefore $\nabla \times \mathbf{A}$, where $\mathbf{A}$ is the difference of the two vector potentials, must be zero throughout the volume v and hence, in general, the field $\mathbf{B} = \nabla \times \mathbf{A}$ is unique. Since it is the curl of $\mathbf{A}$ that is unique, we have defined $\mathbf{A}$ itself only within an additive irrotational vector.

8-7 The use of the vector potential in the solution of problems. We have proved that the tangential component of the magnetic induction field or the tangential component of the vector potential on the surface S uniquely defines the magnetic field within the volume bounded by this surface. This is equivalent to the analogous electrostatic consideration in which the value of the scalar potential on the bounding surface or the value of the normal electrostatic field defines the electrostatic field in the volume bounded by this surface. It is possible to carry the analogy of this procedure still further by writing the vector potential within v explicitly in terms of the currents $\mathbf{j}$ within v and the boundary values of the field over S which are chosen such as to make the field outside S equal to zero. The surface terms will then correspond to the complementary solution of the differential equation, while the volume integral of the currents will correspond to the particular integral. We shall not carry out the details of this process here.* The particular integral is just that which becomes equal to the general solution given by Eq. (1-5) or, more specifically, by Eq. (7-39), in case we let the boundary be at infinity, i.e., the solution that corresponds to knowing the sources over all space.

The differential equation for $\mathbf{A}$ must be derived from the field equations by use of a vector identity for the double curl. This gives

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2\mathbf{A} = \mu_0\mathbf{j}. \tag{8-25}$$

Some care must be used in the interpretation of the operation of the symbol ∇^2 when it is applied to a vector. In the Cartesian coordinate system $\nabla^2\mathbf{A}$ means a vector whose αth component is $\nabla^2 A_\alpha$. In a non-Cartesian coordinate system it is a vector whose αth component must be evaluated by means of the identity

$$\nabla^2\mathbf{A} = -\nabla \times (\nabla \times \mathbf{A}) + \nabla(\nabla \cdot \mathbf{A}) \tag{8-26}$$

The choice of $\nabla \cdot \mathbf{A}$ has thus far been left arbitrary, since $\mathbf{A}$ was defined only in terms of the equation $\mathbf{B} = \nabla \times \mathbf{A}$. It is convenient here to take

$$\nabla \cdot \mathbf{A} = 0, \tag{8-27}$$

which involves no new physical assumptions. We shall find later in considering nonstationary currents that a somewhat more complicated expression must be substituted for Eq. (8-27) in order to preserve symmetry between the electric and magnetic cases, and in the more general application to obtain relativistic covariance of the resulting equations. There will be no conflict with Eq. (8–27) for stationary currents, however.

---

* *Cf.* J. A. Stratton, *Electromagnetic Theory*, p. 245 ff.

With the simplifying assumption of Eq. (8–27) the differential equation for **A** reduces to

$$\nabla^2 \mathbf{A} = -\mu_0 \mathbf{j}. \tag{8–28}$$

This is the vector form of Poisson's equation, of which the particular integral is

$$\mathbf{A} = \frac{\mu_0}{4\pi} \int \frac{\mathbf{j}}{r}\, dv', \tag{8–29}$$

which becomes the general solution if the integral extends over all of the currents that contribute to the field. This is identical to Eq. (7–39); the assumption $\nabla \cdot \mathbf{A} = 0$ corresponds to setting the arbitrary function generated by integrating Eq. (7–38) equal to zero.

The solution of Eq. (8–28) subject to arbitrary boundary conditions is usually considerably more complicated than that of the corresponding scalar potential equation. The reason for this is that, because of the restriction of Eq. (8–27), **A** does not actually have three independent components. This means that we cannot expect to expand the components of **A** separately in normal orthogonal functions and have a sufficient number of boundary conditions to determine the coefficients. In other words, each "harmonic" of the separated scalar solution involving a single coordinate includes two constants of integration, and thus there are six constants in all for each term of the series. Here there are not eighteen constants, six for each component of the vector, but, because of the condition div $\mathbf{A} = 0$, only twelve constants altogether.

We therefore seek an expression for **A** which involves two scalar potential functions, say U and V, and which reduces the equation $\nabla \cdot \mathbf{A} = 0$ to an identity. Let us examine a solution of the type

$$\mathbf{A} = \nabla V + \nabla \times (\mathbf{a}_1 U), \tag{8–30}$$

which obviously satisfies $\nabla \cdot \mathbf{A} = 0$. Here $\mathbf{a}_1$ is one of the unit vectors of the (in general curvilinear) coordinate system. The expression $\mathbf{B} = \nabla \times \mathbf{A}$ will not depend on V; hence U is the only function necessary to specify the field. This is as it should be, because in a source-free region there is no distinction between a field describable by a scalar or a vector potential. In general, however, V is needed to meet boundary requirements on **A**.

The function **A** as given by Eq. (8–30) above satisfies the condition $\nabla \times (\nabla \times \mathbf{A}) = 0$ if the unit vector $\mathbf{a}_1$ is such that $\nabla \times \mathbf{a}_1 = 0$; this precludes $\mathbf{a}_1$ from corresponding to angular variables, although it may be taken as the unit vector along axial or radial variables. Under this restriction, we have

$$\begin{aligned} \mathbf{B} = \nabla \times \mathbf{A} &= -\nabla \times (\mathbf{a}_1 \times \nabla U) \\ &= -\mathbf{a}_1 \nabla^2 U + \nabla[(\mathbf{a}_1 \cdot \nabla) U]. \end{aligned}$$

Since U is harmonic, the equation $\nabla \times (\nabla \times \mathbf{A}) = 0$ is satisfied, and thus $\mathbf{B}$ is derivable from a scalar function. The second part of vector $\mathbf{A}$ as defined by Eq. (8–30) has no component along $\mathbf{a}_1$, since $\mathbf{a}_1 \cdot [\nabla \times (\mathbf{a}_1 U)] = \nabla \cdot [\mathbf{a}_1 \times (\mathbf{a}_1 U)] = 0$ if $\nabla \times \mathbf{a}_1 = 0$.

The boundary conditions on $\mathbf{A}$ at a magnetic interface are derived from the conditions on $\mathbf{B}$. Conservation of magnetic flux across the boundary demands conservation of the line integral of $\mathbf{A}$ along any arbitrary curve in the boundary surface. We thus require that on the interface between regions 1 and 2

$$\mathbf{A}_{t_1} = \mathbf{A}_{t_2},$$

where the subscript t denotes the tangential component. In the absence of true surface currents the tangential component of $\mathbf{H}$ is conserved; if the permeabilities are linear this is equivalent to

$$\frac{1}{\mu_1} (\nabla \times \mathbf{A})_{t_1} = \frac{1}{\mu_2} (\nabla \times \mathbf{A})_{t_2}.$$

Both of these boundary conditions have two components, and hence we have four boundary equations. Note that in the analogous scalar case we had two conditions, Eqs. (2–15) and (2–19), on the scalar potential ϕ. Here we have a sufficient number of conditions to join V and U of Eq. (8–30) across a boundary.

In practical cases not involving boundaries it is conventional to write down solutions of Eq. (8–30) with $V = 0$; then, for a given choice of $\mathbf{a}_1$, orthogonal expansions for $\mathbf{A}$ can be made by using the same number of functions as in the scalar case.

The derivation of $\mathbf{A}$ from scalar solutions of Laplace's equation also serves to circumvent an additional complication arising from the fact that only in rectangular coordinates is $\nabla^2 \mathbf{A}$ simply the vector sum of the Laplacians of the separate components of $\mathbf{A}$. We shall meet this problem again in connection with the vector wave equation.

8–8 The vector potential in two dimensions. In two-dimensional problems, where it can be assumed that the fields are not functions of the z-coordinate, a simple use can be made of the vector potential. In the case in which all lines of flow are parallel to the z-axis, it follows from Eq.

(8–29) that the vector potential has only a z-component. The magnetic fields are then derived from the potential by

$$B_x = \frac{\partial A_z}{\partial y}, \quad B_y = -\frac{\partial A_z}{\partial x}, \tag{8–31}$$

and Laplace's equation becomes

$$\nabla^2 A_z = 0. \tag{8–32}$$

Since A_z is not a function of z, the divergence of $\mathbf{A}$ is obviously equal to zero. Equation (8–32) is the two-dimensional Laplace's equation, and thus we can make the vector potential either the real or the imaginary part of a complex potential. Equations (8–31) are mathematically the same as the equations that relate the stream function to the corresponding electrostatic field. Since the form of the potential for each rectangular component of the vector potential is the same as the scalar Coulomb potential, the vector potential of the line current in two dimensions has the form

$$A_z = \frac{\mu_0}{2\pi} J \ln\left(\frac{r}{r_1}\right), \tag{8–33}$$

and the corresponding complex potential is given by

$$W = \phi + i\psi = \frac{i\mu_0}{2\pi} J \ln\left(\frac{x + iy}{x_1 + iy_1}\right). \tag{8–34}$$

The imaginary part of W, the stream function ψ, is the vector potential. All the methods developed for finding the stream function in the solution of electrostatic problems in two dimensions can be used for the solution of two-dimensional magnetic boundary value problems.

Complex transformations, such as the Schwarz transformation, can be used to transform simple problems involving currents in the neighborhood of permeable media having rectilinear boundaries into more complicated configurations. For example, the problem of a line current located at a given distance from the surface of a semi-infinite permeable medium is soluble by the method of images (the image current here is of the same magnitude and sign as the current in the original conductor) and therefore the solution of various problems concerning slots or gaps in permeable materials under the influence of magnetizing windings can be derived from this simple image solution by means of a suitable Schwarz transformation.

The method of two-dimensional harmonics is also useful, particularly in problems involving cylinders of permeable material. Consider, for example, a permeable conducting cylinder of radius a, carrying a steady current J in the presence of an external field $\mathbf{B}_0$ at right angles to its axis. As usual, we may divide the plane into two regions. For $r > a$ the vector potential A_z is a solution of Laplace's equation, which for large r must give the constant field as well as that produced by the current J. For $r < a$ we must solve the inhomogeneous equation

$$\nabla^2 A_z = -\mu_i j, \tag{8-35}$$

where $j = J/\pi a^2$ and μ_i is the permeability of the cylinder. A particular solution of Eq. (8-35) is $-\mu_i j r^2/4$, and since negative powers of r are excluded by the condition that the origin be a regular point, the general solution is

$$\underset{r<a}{A_z} = -\frac{\mu_i j r^2}{4} + \sum_{n=0}^{\infty} (a_n \cos n\varphi + b_n \sin n\varphi) r^n. \tag{8-36}$$

In order to write a solution valid for the region outside the cylinder, we note that a constant field in the x-direction may be derived from the vector potential $B_0 y = B_0 r \sin \varphi$, and that the effect of the current is that of a line current at the origin. Therefore

$$\underset{r>a}{A_z} = \frac{\mu_0 J}{2\pi} \ln \frac{a}{r} + B_0 r \sin \varphi + \sum_{n=0}^{\infty} (c_n \cos n\varphi + b_n \sin n\varphi) r^{-n}. \tag{8-37}$$

The continuity of the normal component of **B** at the boundary $r = a$ is equivalent to the continuity of A_z, while the continuity of the tangential component of **H** demands that

$$\frac{1}{\mu_i} \frac{\partial}{\partial r} (A_{z<a}) = \frac{1}{\mu_0} \frac{\partial}{\partial r} (A_{z>a}) \tag{8-38}$$

at $r = a$. These two conditions serve to determine the coefficients, with the result that

$$\begin{aligned} \underset{r<a}{A_z} &= -\frac{\mu_i j r^2}{4} + \frac{2\mu_i}{\mu_i + \mu_0} B_0 r \sin \varphi, \\ \underset{r>a}{A_z} &= \frac{\mu_0 J}{2\pi} \ln \left(\frac{a}{r}\right) + \left(1 + \frac{\mu_i - \mu_0}{\mu_i + \mu_0} \frac{a^2}{r^2}\right) B_0 r \sin \varphi, \end{aligned} \tag{8-39}$$

from which B_r and **B** may be immediately computed. It is instructive to write the potential valid inside the cylinder in rectangular coordinates:

$$\underset{r<a}{A_z} = -\frac{\mu_i j (x^2 + y^2)}{4} + \frac{2\mu_i}{\mu_i + \mu_0} B_0 y, \tag{8-40}$$

from which it is more easily seen that magnetization produced by the external field is uniform, and that in the limit of large permeability the induction field is twice that applied, in addition to that produced by the current.

8–9 The vector potential in cylindrical coordinates. A three-dimensional case of practical importance is that of cylindrical symmetry, i.e., current flow in coaxial circles only. For this case the only component of **A** is A_φ, which satisfies the differential equation

$$[\nabla \times (\nabla \times \mathbf{A})]_\varphi = \frac{\partial^2 A_\varphi}{\partial r^2} + \frac{1}{r}\frac{\partial A_\varphi}{\partial r} - \frac{A_\varphi}{r} + \frac{\partial^2 A_\varphi}{\partial z^2} = -\mu j_\varphi \qquad (8\text{–}41)$$

in cylindrical coordinates r, φ, z, since the solution is independent of the coordinate φ. As indicated in Section 5–8, the solutions are of the form

$$A_\varphi = \cos(kz - \alpha_k)[A(k)I_1(kr) + B(k)K_1(kr)], \qquad (8\text{–}42)$$

or

$$A_\varphi = e^{\mp kz}[A^\pm(k)J_1(kr) + B^\pm(k)N_1(kr)], \qquad (8\text{–}43)$$

according to whether real or imaginary values of the separation parameter k are appropriate. If the current flows in a loop or solenoid of radius r_0 the solution may be written in the form of Eq. (8–42) for region 1, $r < r_0$, and for region 2, $r > r_0$. The constants may be determined by the use of the boundary conditions on the interface, $r = r_0$, between regions 1 and 2:

$$A_\varphi^{(1)} = A_\varphi^{(2)}, \qquad (8\text{–}44)$$

$$\frac{1}{\mu_1}\frac{1}{r}\frac{\partial}{\partial r}(rA_\varphi^{(1)}) - \frac{1}{\mu_0}\frac{1}{r}\frac{\partial}{\partial r}(rA_\varphi^{(2)}) = j_{s\varphi}, \qquad (8\text{–}45)$$

where $j_{s\varphi}$ is the surface current density at the interface. Equations (8–44) and (8–45) correspond to $\nabla \cdot \mathbf{B} = 0$ and $\nabla \times \mathbf{H} = \mathbf{j}$ respectively. (The z-component of **H** is discontinuous at $r = r_0$.) If the current loops are in a plane, which we may take as $z = 0$, the solution should be written in the form of Eq. (8–43) for the regions of positive and negative z. The boundary condition replacing Eq. (8–45) is that the radial component of **H**, namely $(\partial A_\varphi/\partial z)/\mu$, suffers a discontinuity corresponding to the surface current in the plane $z = 0$.

Let us apply this method to the determination of the vector potential due to a current J in a single plane loop of radius r_0. In this case the choice of functions is optional. If we choose Eq. (8–42) so as to take the cylinder of radius r_0 as the surface dividing the two regions, we first note that $K_1(kr)$ must be excluded from the inner region and $I_1(kr)$ from the outer region in order that our solutions be regular everywhere. Therefore we may write

$$\begin{aligned} A_\varphi^{(1)} &= \int_0^\infty A(k)I_1(kr)\cos kz\, dk, \\ A_\varphi^{(2)} &= \int_0^\infty B(k)K_1(kr)\cos kz\, dk. \end{aligned} \qquad (8\text{–}46)$$

Equation (8–44) demands that

$$A(k)I_1(kr_0) = B(k)K_1(kr_0), \tag{8–47}$$

while Eq. (8–45), with $j_\varphi = J\,\delta(z)$, gives

$$\left[\int_0^\infty \left\{A(k)\,\frac{\partial}{\partial r}\,[rI_1(kr)] - B(k)\,\frac{\partial}{\partial r}\,[rK_1(kr)]\right\}\frac{\cos kz\,dk}{r}\right]_{r=r_0} = \mu J\,\delta(z). \tag{8–48}$$

We can multiply by $\cos k'z$ and apply the Fourier integral theorem or, equivalently, remember the δ-function character of $\int \cos kz\,dk$, to obtain from Eq. (8–48)

$$\frac{\pi}{r_0}\left[A(k)\,\frac{\partial}{\partial r}\,[rI_1(kr)] - B(k)\,\frac{\partial}{\partial r}\,[rK_1(kr)]\right]_{r=r_0} = \mu J. \tag{8–49}$$

But there is a mathematical relation between the Bessel functions and their derivatives, namely,

$$I_n'(kr)K_n(kr) - K_n'(kr)I_n(kr) = \frac{1}{kr}\cdot \tag{8–50}$$

Therefore we may obtain the coefficients at once,

$$A(k) = \frac{r_0\mu J}{\pi}\,K_1(kr_0),$$

$$B(k) = \frac{r_0\mu J}{\pi}\,I_1(kr_0),$$

and the potentials are

$$\begin{aligned} A_\varphi^{(1)} &= \frac{r_0\mu J}{\pi}\int_0^\infty K_1(kr_0)I_1(kr)\cos kz\,dk, \\ A_\varphi^{(2)} &= \frac{r_0\mu J}{\pi}\int_0^\infty I_1(kr_0)K_1(kr)\cos kz\,dk. \end{aligned} \tag{8–51}$$

If, on the other hand, the $z = 0$ plane is used as the division between the two regions, Eq. (8–43) leads to

$$A_\varphi^\pm = \int A^\pm(k)J_1(kr)e^{\mp kz}\,dk, \tag{8–52}$$

since $N_1(kr)$ is not regular at the origin. The surface current is confined to $J\,\delta(r - r_0)$ in the plane, and introduces a discontinuity in $\partial A_\varphi/\partial z$ at

$z = 0$. The determination of the coefficients $A(k)$ requires the use of the Fourier-Bessel integral

$$\int_0^\infty k\,dk \int_0^\infty r\,dr\, f(k) J_n(kr) J_n(k'r) = f(k'), \tag{8-53}$$

which leads at once to

$$A_\varphi^\pm = \frac{r_0 \mu J}{2} \int_0^\infty J_1(kr_0) J_1(kr) e^{\mp kz}\, dk \tag{8-54}$$

for the required potential.

SUGGESTED REFERENCES

J. A. STRATTON, *Electromagnetic Theory.* Chapter IV includes a derivation of the complete solution for the vector potential including boundary conditions.

P. M. MORSE AND H. FESHBACH, *Methods of Theoretical Physics.* A more complete mathematical treatment than that of Stratton.

W. R. SMYTHE, *Static and Dynamic Electricity.* Chapter VII contains methods for determining the vector potential and many examples of its use.

EXERCISES

1. Consider a magnet with pole pieces wide compared with the pole gap, and with the windings far removed from the gap. Find the complex potential function and plot the field in the plane of symmetry from a point well outside the poles to a position in the gap where the field is essentially constant.

2. Find **B** and **H** inside and outside a spherical shell of radii a and b which is magnetized permanently to a constant magnetization **M**. What is the effect of making the spherical cavity not concentric with the outside surface of the shell?

3. A cylindrical hole of radius a is bored parallel to the axis of a long cylindrical conductor of radius b which carries a uniformly distributed current of density $\mathbf{j}$. The distance between the center of the conductor and the center of the hole is x_0. Find the magnetic field in the hole. (The generalization of this result to cylinders of elliptical cross section is used in the design of the magnetic field for some high-energy accelerating devices.)

4. A coil is wound on the surface of a sphere such that the field inside the sphere is uniform. What is the winding? (This form of winding is used in the Westinghouse-Goudsmit mass spectrometer.)

5. What is the effective dipole moment of a sphere at points outside if the sphere has a uniform surface charge density and rotates with angular velocity ω?

6. A spherical shell of material with permeability $\mu \neq \mu_0$, whose inner and outer radii are a and b respectively, is placed in an originally uniform field B_0. What is the field in the spherical cavity of radius a?

7. Consider an infinitely long wire at $x = a$, $y = 0$, carrying current J in vacuum, while all space for which x is negative is filled with a medium of $\mu \neq \mu_0$. Find the field at all points in space.

8. Find the magnetic field of a long wire at $x = a$, $y = b$, carrying current J in vacuum, while all of space for which x is negative and all for which y is negative are filled with magnetic material of very large permeability, $\mu \to \infty$.

9. Calculate the magnetic vector potential due to a ring of radius a coaxial with a permeable circular cylinder of permeability $\mu \neq \mu_0$ and radius $b < a$ (Smythe).

10. A soft iron ring of radius b, cross section radius $a \ll b$, is wound with N turns of wire carrying current J. A small air gap of length δ is cut at one point in the ring. Assuming a large permeability for the iron, compute the reluctance of the magnetic circuit.

11. Show that if $\mathbf{B} = \nabla \times \mathbf{A}$ and $\mathbf{H} = -\nabla\phi_m$, and if $\mathbf{A}$ and ϕ_m are given by Eqs. (8–13) and (8–14) respectively, then $\mathbf{B}$ and $\mathbf{H}$ satisfy the relation $\mathbf{B} = \mu_0\mathbf{H} + \mathbf{M}$.

12. Put $\mathbf{a}_1 = \mathbf{k}$ in Eq. (8–30) where $\mathbf{k}$ is the unit vector along the z-axis in cylindrical coordinates. Take $V = 0$. Show that the resulting vector potential expansion, if U is expanded as a scalar potential solution in cylindrical coordinates, agrees with Eqs. (8–42) and (8–43), obtained by direct integration.

CHAPTER 9

MAXWELL'S EQUATIONS

9–1 Faraday's law of induction. In electrostatics the electric field is conservative, i.e., curl $\mathbf{E} = 0$. We have seen in Chapter 7 that in order to produce stationary currents there must be electric fields, such as $\mathbf{E}'$ in Eq. (7–5), which violate this condition. And it is found experimentally that a nonconservative electric field actually accompanies varying magnetic fields. The law describing this situation is usually known as the Faraday law of induction, and can be formulated as follows. Consider a circuit of resistance R carrying a current J and containing an electromotive force $\mathcal{E}$. The magnetic flux linking this circuit is defined by

$$\Phi_m = \int \mathbf{B} \cdot d\mathbf{S}, \tag{9–1}$$

where the surface of integration is bounded by the circuit. If this flux changes in time, it is found experimentally that

$$JR - \mathcal{E} = -\frac{d\Phi_m}{dt}. \tag{9–2}$$

This means that the current in the circuit differs from that predicted by Ohm's law by an amount which can be attributed to an additional electromotive force equal to the negative time rate of change of flux through the circuit. Note that Eq. (9–2) is an independent experimental law and is in no way derivable from any of the relations that have been previously used. In particular, contrary to the statement that is sometimes made, Faraday's law of induction is not the consequence of the law of conservation of energy applied to the over-all energy balance of currents in magnetic fields.

Equation (9–2) is formulated in terms of the total flux passing through the given circuit. This flux can change for several reasons: it can change because of changes in the external field with time; it can change because of motion of the circuit itself or parts of the circuit. We shall consider Eq. (9–2) to be an experimental law which holds for all such cases, including that of currents in moving media.

It was recognized by Maxwell that the Faraday law of induction had a very much more general significance than the case actually described by Eq. (9–2) would indicate. The equation can be written in the equivalent form,

$$\oint \mathbf{E}' \cdot d\mathbf{l} = -\frac{d\Phi_m}{dt} = JR - \mathcal{E}, \tag{9–3}$$

which indicates that there must be an electric field along the wire that is nonelectrostatic. However, from the boundary condition that requires the tangential components of the electric field across the boundary of a wire to be continuous, we can conclude that Eq. (9–3) is also valid in the region immediately adjacent to the wire. Since the characteristics of the wire, namely, its resistance and its electromotive forces, are not contained in the left side of Eq. (9–3), it appears likely that this relation is, in fact, independent of the presence of a current-carrying conductor, and is a general physical law relating an electric field *in vacuo* to the rate of change of a magnetic field. If this is so, then the equation can be transformed into a differential expression valid for free space or a stationary medium. In either case, the total derivative of the flux integral, Eq. (9–1), can be written as an integral of the partial time derivative of the magnetic field, and Eq. (9–3) becomes

$$\oint \mathbf{E}' \cdot d\mathbf{l} = -\frac{d}{dt}\int \mathbf{B} \cdot d\mathbf{S} = -\int \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S}. \tag{9–4}$$

The use of Stokes' theorem leads to

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}. \tag{9–5}$$

Equation (9–5) expresses the modification of the electrostatic field irrotationality which is a necessary consequence of Faraday's law of induction.

9–2 Maxwell's equations for stationary media. We now have expressions for both the source densities (divergence) and circulation densities (curl) of the two basic field vectors **E** and **B**. From Eqs. (2–5), (7–35), (8–3), and (9–5), we have

$$\begin{aligned}
&(1) \qquad \nabla \cdot \mathbf{E} = \epsilon_0^{-1}\rho_{\text{total}} = \epsilon_0^{-1}(\rho_{\text{true}} - \nabla \cdot \mathbf{P}),\\
&(2) \qquad \nabla \cdot \mathbf{B} = 0,\\
&(3) \qquad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \qquad\qquad (9\text{–}6)\\
&(4) \qquad \nabla \times \mathbf{B} = \mu_0\left(\mathbf{j}_{\text{true}} + \frac{\partial \mathbf{P}}{\partial t} + \nabla \times \mathbf{M} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}\right).
\end{aligned}$$

Equations (9–6) are Maxwell's electrodynamic field equations, formulated so as to be valid for media at rest. The restriction to material media at rest arises from the omission of any convective current terms from Eq. (9–6) (4), and from ignoring, in the derivation of Eq. (9–5), flux changes due to motion of the medium. Equations (9–6) are written in terms of the equivalent vacuum charges or currents which give rise to the fields, and contain the equivalent current and charge densities explicitly. If the additional field vectors **D** and **H** are introduced by the defining equations

$$\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P}, \tag{2–3}$$

$$\mathbf{H} = \frac{\mathbf{B}}{\mu_0} - \mathbf{M}, \tag{8–4}$$

Maxwell's field equations become

$$\begin{aligned}
&(1) \qquad \nabla \cdot \mathbf{D} = \rho_{\text{true}},\\
&(2) \qquad \nabla \cdot \mathbf{B} = 0,\\
&(3) \qquad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t},\\
&(4) \qquad \nabla \times \mathbf{H} = \mathbf{j}_{\text{true}} + \frac{\partial \mathbf{D}}{\partial t}.
\end{aligned} \tag{9–7}$$

The field equations (9–7) appear formally simpler than Eqs. (9–6), but they are actually more complicated physically. The solution of these equations is possible only if additional constitutive equations are available connecting **D** to **E**, **j** to **E**, and **H** to **B**, such as $\mathbf{D} = \kappa\epsilon_0\mathbf{E}$, $\mathbf{j}_{\text{true}} = \sigma\mathbf{E}$, $\mathbf{H} = \mathbf{B}/\mu$, for a linear medium, or whatever forms apply for a nonlinear medium.

9–3 Faraday's law for moving media. It is necessary to use considerable care in extending the law of induction to take account of motion in general. We must first derive the subsidiary theorem which expresses the total time rate of change of the flux across a given surface in terms of a surface integral of the vector function **B**, even when the surface itself across which the flux is evaluated is in motion.

Again let Φ_m be the flux of the vector field **B** across the surface S. We are seeking the function $D\mathbf{B}/Dt$, defined by

$$\frac{d}{dt}\Phi_m = \frac{d}{dt}\int \mathbf{B} \cdot d\mathbf{S} = \int \frac{D\mathbf{B}}{Dt} \cdot d\mathbf{S}. \tag{9–8}$$

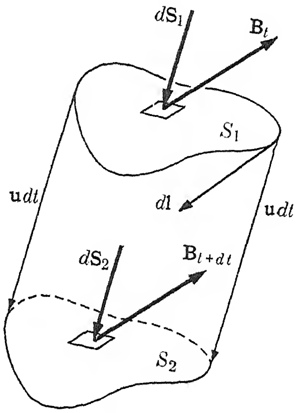


FIG. 9–1 Figure for evaluation of $D\mathbf{B}/Dt$.

In order to evaluate $D\mathbf{B}/Dt$, let us consider the surface in Fig. 9-1 in position 1 at time t, and in position 2 at time $t + dt$. By the rules for differentiation,

$$\frac{\Delta}{\Delta t}\int \mathbf{B}\cdot d\mathbf{S} = \frac{1}{\Delta t}\int (\mathbf{B}_{t+dt}\cdot d\mathbf{S}_2 - \mathbf{B}_t\cdot d\mathbf{S}_1). \tag{9-9}$$

If we apply Gauss's theorem at the time t to the volume enclosed by S_1, S_2, and the traces of the edges of S, we have

$$\int \nabla\cdot\mathbf{B}\,dv = \int (\mathbf{B}_t\cdot d\mathbf{S}_2 - \mathbf{B}_t\cdot d\mathbf{S}_1) - \oint \mathbf{B}_t\cdot(\mathbf{u}\,dt\times d\mathbf{l}). \tag{9-10}$$

The last term represents the flux change across the side surface generated by the motion of the boundary of S, of which an element is $d\mathbf{l}$. Note that the flux across the surfaces S_1 and S_2 in Eq. (9-10) is considered at time t, since Gauss's theorem applied only to instantaneous values of the vector field $\mathbf{B}$. The value of $\mathbf{B}$ on S_2 at time $t + dt$ may be found in terms of its value at t by Taylor's theorem,

$$\mathbf{B}_{t+dt} = \mathbf{B}_t + \frac{\partial\mathbf{B}}{\partial t}dt + \cdots. \tag{9-11}$$

With the substitution of Eqs. (9-10) and (9-11), Eq. (9-9) becomes, in the limit,

$$\frac{d}{dt}\int \mathbf{B}\cdot d\mathbf{S} = \int \frac{\partial\mathbf{B}}{\partial t}\cdot d\mathbf{S} + \oint \mathbf{B}\times\mathbf{u}\cdot d\mathbf{l} + \int \frac{\nabla\cdot\mathbf{B}\,dv}{dt}. \tag{9-12}$$

Then, by using Stokes' theorem and the fact that

$$dv = \mathbf{u}\cdot d\mathbf{S}\,dt, \tag{9-13}$$

we obtain the desired relation,

$$\frac{D\mathbf{B}}{Dt} = \frac{\partial\mathbf{B}}{\partial t} + \nabla\times(\mathbf{B}\times\mathbf{u}) + (\nabla\cdot\mathbf{B})\mathbf{u}. \tag{9-14}$$

The first term of this expression represents the change in the flux through S that is caused by the time variation of the vector field. The second term represents the flux loss across the boundary of the moving surface. The third term arises from the passage of surface S through an inhomogeneous vector field in which flux lines are generated.

Equation (9-14) can now be used to express Faraday's law, Eq. (9-2), in differential form in a moving medium. Since $\mathbf{B}$ is always solenoidal, we have from Eqs. (9–3) and (9–14),

$$\oint \mathbf{E}' \cdot d\mathbf{l} = -\frac{d\Phi_m}{dt} = -\int \left[\frac{\partial \mathbf{B}}{\partial t} + \nabla \times (\mathbf{B} \times \mathbf{u}) \right] \cdot d\mathbf{S}. \qquad (9\text{–}15)$$

We have designated the field by $\mathbf{E}'$ around the circuit, since $\mathbf{E}'$ is to be measured in the moving frame of reference, i.e., Faraday's law applies specifically to the current measured in the wire through which the flux is changing, no matter what might be the cause of the flux change. By applying Stokes' theorem to Eq. (9–15), we arrive at

$$\nabla \times \mathbf{E}' = -\frac{\partial \mathbf{B}}{\partial t} - \nabla \times (\mathbf{B} \times \mathbf{u}), \qquad (9\text{–}16)$$

where $\mathbf{E}'$ still represents the field measured in the moving medium. Equation (9–16) can be written in the form

$$\nabla \times (\mathbf{E}' - \mathbf{u} \times \mathbf{B}) = -\frac{\partial \mathbf{B}}{\partial t}. \qquad (9\text{–}17)$$

But the argument of the curl in Eq. (9–17), $\mathbf{E}' - \mathbf{u} \times \mathbf{B}$, actually represents the field that is measured by a stationary observer. The reason for this is that an observer carrying a charge q through a magnetic field $\mathbf{B}$ with a velocity $\mathbf{u}$ will experience a force, $q(\mathbf{u} \times \mathbf{B})$, in addition to the force of the electric field $\mathbf{E}$ which may also be present. Hence the electric field observed by a stationary observer is equal to the electric field $\mathbf{E}'$ seen by the moving observer minus the effective field $\mathbf{u} \times \mathbf{B}$, i.e., $\mathbf{E} = \mathbf{E}' - \mathbf{u} \times \mathbf{B}$. Hence for a stationary observer, Eq. (9–17) becomes

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}.$$

This means that the differential formulation of Faraday's law of induction is independent of the motion of the medium inside the field. This is as it should be, since Eq. (9–5) is purely a field relation in terms of the equivalent vacuum fields $\mathbf{B}$ and $\mathbf{E}$, and should therefore be independent of the characteristics of the medium, including its motion. However, the electric field observed by the moving observer does contain two terms, namely, the "induced field" produced by the time rate of change of the external magnetic fields, and the "motional field," $\mathbf{u} \times \mathbf{B}$, produced by the motion of the observer in the magnetic field. Note that in this discussion it has been assumed that the electric field proper is not affected by the state of motion of the observer. This assumption is actually justified only in case the motion of the observer is small compared with the velocity of light, as we shall see later. Therefore all the conclusions which we now draw concerning Maxwell's equations in moving media can be applied with confidence only when the velocities of such media are small compared with the velocity of light.

9-4 Maxwell's equations for moving media. We have concluded that the third of Maxwell's four equations in the form of Eq. (9-6) is not affected by the motion of the medium in which the fields are measured. The first and second equations are not affected by this motion, either, since nonrelativistically the charge density of the medium is not affected by the state of motion of the observer. The only modification is a correction to the current in the fourth of Eqs. (9-6), a convective term, and correction to the polarization current. The convection current, due to the motion of the charge density and equivalent polarization charge, is simply $\mathbf{u}(\rho_{\text{true}} - \nabla \cdot \mathbf{P})$. The polarization current is properly given by $D\mathbf{P}/Dt$ instead of $\partial \mathbf{P}/\partial t$, to take account of the charges lost due to the change of polarization flux across the moving surface. The total current which gives rise to the magnetic field is thus composed of the following parts:

(1) True currents, $\mathbf{j}$.

(2) Convective currents, $\mathbf{u}(\rho_{\text{true}} - \nabla \cdot \mathbf{P})$.

(3) Currents caused by the rate of change of the polarization and the motion of polarized media, in analogy with Eq. (9-14), given by

$$\frac{D\mathbf{P}}{Dt} = \frac{\partial \mathbf{P}}{\partial t} + \nabla \times (\mathbf{P} \times \mathbf{u}) + (\nabla \cdot \mathbf{P})\mathbf{u}.$$

(4) Vacuum displacement current, $\epsilon_0\, \partial \mathbf{E}/\partial t$.

With these corrections to the fourth of Maxwell's equations, it is now possible to write the entire set of field equations so as to be valid in a nonmagnetized medium moving with a velocity $\mathbf{u}$ which is small compared with the velocity of light:

$$(1) \qquad \nabla \cdot \mathbf{D} = \rho_{\text{true}},$$

$$(2) \qquad \nabla \cdot \mathbf{B} = 0,$$

$$(3) \qquad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t},$$

$$\begin{aligned} (4) \qquad \nabla \times \mathbf{B} &= \mu_0 \left[\mathbf{j}_{\text{true}} + \mathbf{u}(\rho_{\text{true}} - \nabla \cdot \mathbf{P}) + \frac{\partial \mathbf{P}}{\partial t} + \nabla \times (\mathbf{P} \times \mathbf{u}) \right. \\ &\qquad \left. + (\nabla \cdot \mathbf{P})\mathbf{u} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right] \\ &= \mu_0 \left[\mathbf{j}_{\text{true}} + \rho_{\text{true}}\mathbf{u} + \frac{\partial \mathbf{D}}{\partial t} + \nabla \times (\mathbf{P} \times \mathbf{u}) \right]. \end{aligned} \tag{9-18}$$

The constitutive equations which give the true currents in the moving medium and the polarization of the moving medium are derived from the fields measured in the moving medium:

$$\begin{aligned} \mathbf{j} &= \sigma(\mathbf{E} + \mathbf{u} \times \mathbf{B}) = \sigma \mathbf{E}', \\ \mathbf{P} &= \epsilon_0(\kappa - 1)(\mathbf{E} + \mathbf{u} \times \mathbf{B}). \end{aligned} \tag{9-19}$$

For a noncharged and current free dielectric, Maxwell's equations can then be written in the form:

$$\begin{aligned} &(1) & \nabla \cdot \mathbf{D} &= 0, \\ &(2) & \nabla \cdot \mathbf{B} &= 0, \\ &(3) & \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, \\ &(4) & \nabla \times [\mathbf{B} - \mu_0 \mathbf{P} \times \mathbf{u}] &= \mu_0 \frac{\partial \mathbf{D}}{\partial t}. \end{aligned} \tag{9-20}$$

It is seen from (4) of Eqs. (9–20) that from the macroscopic point of view a moving polarized dielectric is equivalent to a magnetized material of magnetic moment

$$\mathbf{M}_{\text{eq}} = \mathbf{P} \times \mathbf{u}. \tag{9-21}$$

This can be easily understood by considering a polarized slab of material moving at right angles to the direction of polarization. In these circumstances, there is an equivalent positive current in the direction of motion, and another, due to the motion of the negative charges and displaced from the first, in the opposite direction. These currents constitute a net current loop, and thus give rise to a magnetic moment. Hence the moving polarized dielectric will produce a magnetic field which is indistinguishable from that of a magnetized material. This effect has been demonstrated experimentally by Roentgen, Eichenwald, and others.

9–5 Motion of a conductor in a magnetic field. As an example of the application of Maxwell's equations to fields in moving media, let us consider a conducting bar, infinitely long and of rectangular cross section, moving in the direction of its length with velocity $\mathbf{u}$ relative to a uniform magnetic field $\mathbf{B}$. As shown in Fig. 9–2, $\mathbf{B}$ is at right angles to the direction of motion of the bar, and we shall assume that it is constant in time. Two sliding contacts, terminals of a stationary galvanometer, touch the bar on opposite sides, as shown, so that the distance l between them is the width of the bar.

From physical considerations, one would expect a current to flow in the external stationary loop. An electron moving with the bar will experience an effective field given by $\mathbf{E}' = \mathbf{u} \times \mathbf{B}$, so that a current will flow through the contacts to be measured in the external circuit by a stationary observer. Curl **E** must vanish for this observer, since **B** is not changing in time. (Whether the source of **B** is stationary or is in motion is entirely irrelevant, since any observed phenomena which depend on a field description must be describable in terms of the behavior of the field quantities alone, independent of the nature of the mechanism which produces the field quantities.) Hence if, as appears logical from the above electron argument, there is an electric field, then such a field must be irrotational, i.e., electrostatic.

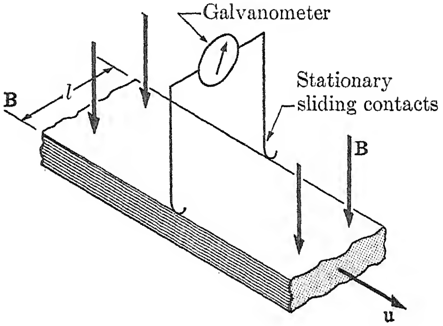


FIG. 9–2 Conducting bar moving along its length in a magnetic field.

The effective electric fields within the moving bar will cause a current within the bar that will produce charges on the faces, and these charges will produce the observed external electrostatic field. On the other hand, the same charge displacement will exactly cancel the effective electric field $\mathbf{u} \times \mathbf{B}$ within the bar, and therefore if we consider an integration path partially contained in the bar and partially outside the bar the integral of the electric field around a closed loop will not vanish. This result is in agreement with the physically observed result that an electromotive force of magnitude uBl is measured across the bar. Note that if the entire circuit were stationary there would be no electromotive force.

If, instead of moving the bar, we move the galvanometer link relative to the bar in the field **B**, again the electromotive force uBl is observed, since the roles of the link and the bar are simply interchanged in the above integration. The various cases of relative motion are summarized in Table 9–1. Two salient facts characterize the results: (1) the state of

TABLE 9–1

ELECTROMOTIVE FORCE RESULTING FROM VARIOUS POSSIBILITIES OF MOTION FOR THE BAR, SOURCE OF MAGNETIC FIELD, AND OBSERVER OF FIG. 9–2

Motion of			Electromotive force measured by the observer
Bar	Source of B	Observer (galvanometer)	
u	0	0	uBl
0	u	0	0
0	u	u	uBl
u	u	0	uBl
0	0	u	uBl
u	0	u	0

motion of the source of **B** is irrelevant as long as **B** is uniform; and (2) absolute motion cannot be detected in this arrangement. The latter fact is an indication that Maxwell's equations, if carefully interpreted, are in agreement with relativistic principles. This will be shown later in greater detail and generality.

The situation is more complicated if, in addition to an external magnetic field, there is a field caused by the magnetic moment, either induced or permanent, of the slab. Our conclusion that the electric field observed in a stationary loop will be purely electrostatic remains valid. However, the source of the electrostatic field will not become fully clear until the equations for moving media are modified to include permeable media. Unfortunately, this modification cannot be made in a reasonable way without introducing relativistic considerations. Nevertheless, the result is physically clear, since the source of the magnetic field, provided it is constant in time, does not affect the force produced by the field. In other words, the force which acts on a moving electron within a moving bar is independent of whether the magnetic field is external or due to the magnetic moment of the bar itself. Therefore we should expect to obtain an electromotive force, given by uBl as before, where **B** is the magnetic field in the moving magnetized bar. The only thing that seems paradoxical is that since a moving magnet is essentially an assembly of current loops, we have apparently concluded that the motion of loops carrying steady currents gives rise to an electrostatic field. This, as will be shown later by relativistic considerations, is in fact true. We shall be able to show in general that if an observer moves with a velocity **u** relative to a medium of magnetization **M** he will observe an equivalent electric moment given by

$$\mathbf{P}_{\text{eq}} = \frac{1}{c^2}\mathbf{u} \times \mathbf{M}. \tag{9-22}$$

Therefore,

$$\nabla \cdot \mathbf{E} = -\frac{\nabla \cdot \mathbf{P}_{\text{eq}}}{\epsilon_0} \tag{9-23}$$

will define the sources of the field. Note that this effect, although it appears deceptively similar to the classical effect of Eq. (9-21), is actually explainable only in terms of the special theory of relativity. It is caused by the fact that to the observer the time spent by the charge traveling in the direction parallel to the relative motion of the circuit and the observer is different from the time that it spends moving in the antiparallel direction. This gives rise to an effective polarization that is perpendicular to the direction of motion and which lies in the plane of the current loop. We shall discuss this effect in detail later.

If the length of the magnetized slab is finite, the field **E** will no longer be irrotational, since $\partial \mathbf{B}/\partial t$ is no longer zero in the rest frame. In fact, since $\partial/\partial t = -\mathbf{u} \cdot \nabla$ for uniform motion $\mathbf{u}$, $\nabla \times \mathbf{E} = -\partial \mathbf{B}/\partial t = (\mathbf{u} \cdot \nabla)\mathbf{B}$, which because of the vanishing divergence of $\mathbf{B}$ reduces to $-\nabla \times (\mathbf{u} \times \mathbf{B})$. Hence if $\mathbf{B}$ is no longer uniform, then $\mathbf{E}$ is no longer irrotational; its curl, however, is identical with the curl of $-(\mathbf{u} \times \mathbf{B})$, which is also the effective electric field in the moving medium.

SUGGESTED REFERENCES

The law of induction and its consequences are treated in all books on electrodynamics. The treatment most nearly parallel to ours is that in *Classical Electricity and Magnetism*, Vol. 1, M. Abraham and R. Becker.

EXERCISES

1. Show that if a flip coil is wound on the surface of a sphere with a winding density proportional to $\sin \theta$, where θ is the polar angle, then this coil will measure the axial field component at the center of the coil, independent of the degree of homogeneity of the field. (See Exercise 4, Chapter 8.)
2. Approximate the properties of the coil of problem 1 by a simple cylindrical coil. What are the optimum proportions (ratio of height to diameter) of such a coil to minimize gradient effects?
3. A dielectric ($\kappa > 1$) cylinder of radius a is oriented with its axis parallel to a magnetic induction field $\mathbf{B}$. If it is rotated about its axis with an angular velocity ω, find the resultant polarization per unit volume, and the charge per unit length that appears on its surface.

CHAPTER 10

ENERGY, FORCE, AND MOMENTUM RELATIONS IN THE ELECTROMAGNETIC FIELD

In the discussion of the energy relations in electrostatic fields (Chapter 6) we succeeded in associating an energy density with the electric field by considering a specific process. In that case the process was the assembly of charges, during which work was done and changes were produced in the fields. It was possible to obtain a free energy density of the electric field, in the thermodynamic sense, by balancing the work and energy terms. Let us consider an analogous process to establish a magnetic field before proceeding to obtain a general set of relations for the electromagnetic field.

10–1 Energy relations in quasi-stationary current systems. Consider a process in which a battery with a nonelectrostatic field $\mathbf{E}'$ is feeding energy both into heat losses and into a magnetic field. If we take the scalar product of $\mathbf{j}$ and the equation

$$\mathbf{j} = \sigma(\mathbf{E} + \mathbf{E}'), \tag{7–4}$$

we obtain

$$\mathbf{E}' \cdot \mathbf{j} = \frac{j^2}{\sigma} - \mathbf{E} \cdot \mathbf{j}. \tag{10–1}$$

The left side of Eq. (10–1) represents the time rate at which the battery does work; the first term on the right represents the Joule heat loss in the current-carrying medium, and the last term we tentatively identify as the rate at which energy is fed into the magnetic field. If all fields are quasi-stationary, i.e., slowly varying, the displacement current may be neglected, and the fourth of Maxwell's equations becomes $\nabla \times \mathbf{H} = \mathbf{j}$. When we make this substitution in Eq. (10–1) and integrate over all space, we obtain an expression for the total power expended by the battery in terms of the fields:

$$\int \mathbf{E}' \cdot (\nabla \times \mathbf{H})\, dv = \int \frac{(\nabla \times \mathbf{H})^2}{\sigma}\, dv - \int \mathbf{E} \cdot (\nabla \times \mathbf{H})\, dv. \tag{10–2}$$

The last term may be integrated by parts by means of the relation

$$\nabla \cdot (\mathbf{E} \times \mathbf{H}) = \mathbf{H} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot (\nabla \times \mathbf{H}), \tag{10–3}$$

and if we use the divergence theorem and the third Maxwell equation, we find that

$$\int \mathbf{E} \cdot (\nabla \times \mathbf{H})\, dv = -\int \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t}\, dv - \int (\mathbf{E} \times \mathbf{H}) \cdot d\mathbf{S}. \qquad (10\text{–}4)$$

Since $\mathbf{E} \times \mathbf{H}$ falls off at least as $1/r^5$ in electrostatic and quasi-stationary magnetic fields, the surface integral in Eq. (10–4) may be dropped. Note that this will not be possible in case $\mathbf{E}$ and $\mathbf{H}$ fall off as $1/r$, as they do in radiation fields. In the absence of the displacement current term no such fields arise: our restriction to slowly varying fields corresponds to neglect of all radiation terms in the magnetic field. Similar limitations apply to our considerations of electrostatic field energy. Thus far we are using, separately, energy relations in electrostatic fields on the one hand, and quasi-stationary current magnetic fields on the other. Later we must see how these concepts can be modified in a consistent way to obtain the general energy expressions.

The total power expended by the battery, Eq. (10–2), may thus be written

$$\int \mathbf{E}' \cdot (\nabla \times \mathbf{H})\, dv = \int \frac{(\nabla \times \mathbf{H})^2}{\sigma}\, dv + \int \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t}\, dv. \qquad (10\text{–}5)$$

The first term on the right has already been identified with the rate of Joule heat loss. The last term is now obviously the rate at which energy is fed into the field. The variation δU_m in magnetic field energy can therefore be given by

$$\delta U_m = \int \mathbf{H} \cdot \delta \mathbf{B}\, dv. \qquad (10\text{–}6)$$

This is analogous to the electrostatic energy variation

$$\delta U = \int \mathbf{E} \cdot \delta \mathbf{D}\, dv. \qquad (6\text{–}19)$$

As was the case with U, U_m represents the free energy.

In order to make Eq. (10–6) integrable, we must assume a functional relation between $\mathbf{H}$ and $\mathbf{B}$. For a medium which magnetizes linearly the integration can be carried out in the same manner as Eq. (6–14), giving

$$U_m = \tfrac{1}{2} \int \mathbf{H} \cdot \mathbf{B}\, dv. \qquad (10\text{–}7)$$

In nonlinear materials such as ferromagnets Eq. (10–6) can be integrated only between definite states, and the answer will, in general, depend on the past history of the sample. For ferromagnets the integral of Eq. (10–6) has a finite value, not zero, when **B** is evaluated around a complete cycle, as in a field produced by an alternating current. The cyclic energy loss is given by

$$\Delta U_m = \int \oint \mathbf{H} \cdot d\mathbf{B}\, dv. \tag{10–8}$$

This means that the energy expended per unit volume when a magnetic material is carried through a magnetization cycle is equal to the area of its hysteresis loop as plotted in a graph of B against H.

Equation (10–7) gives the energy in terms of a volume integral over the fields. If we wish an expression as a volume integral over the sources, we need only write **B** in terms of the vector potential **A**, and **H** in terms of the stationary current **j**. First we obtain

$$U_m = \tfrac{1}{2} \int \mathbf{H} \cdot (\nabla \times \mathbf{A})\, dv. \tag{10–9}$$

If we now integrate by parts and drop the surface terms as before, we find that

$$U_m = \tfrac{1}{2} \int \mathbf{j} \cdot \mathbf{A}\, dv, \tag{10–10}$$

which is analogous to the expression for the electrostatic energy in terms of volume charge density and the scalar potential.

Equations (10–6), (10–7), and (10–10) have been derived by a particular "virtual process." The expressions depend only on the final fields, however, and not on the nature of the process. They may thus be taken to represent the magnetic field energy in general, with Eqs. (10–7) and (10–10) subject only to the restrictions of a linear relation between **B** and **H**.

The factor $\frac{1}{2}$ in Eq. (10–10) is similar to the factor $\frac{1}{2}$ in Eq. (6–1), and is due to the fact that the vector potential **A** includes the fields of the currents **j** themselves. The interaction energy of a system of currents and charges in an *external* field of potentials ϕ and **A** is given by

$$U_{\text{interaction}} = \int (\mathbf{j} \cdot \mathbf{A}_{\text{external}} + \rho\phi_{\text{external}})\, dv. \tag{10–11}$$

10–2 Forces on current systems. We shall now use the energy expressions for two purposes: first, to derive expressions for the forces between currents in terms of the currents themselves and suitable geometrical parameters which depend on their location; and second, to express the variation of the magnetic field energy in terms of the variations of the currents and of the geometrical coordinates.

Let us analyze a system of n geometrically linear circuits carrying currents J_k. For these line circuits Eq. (10–10) reduces to

$$U_m = \tfrac{1}{2}\sum_{k=1}^{n} J_k \oint \mathbf{A}\cdot d\mathbf{l}_k = \tfrac{1}{2}\sum_{k=1}^{n} J_k \int (\nabla\times\mathbf{A})\cdot d\mathbf{S}_k$$

$$= \tfrac{1}{2}\sum_{k=1}^{n} J_k \int \mathbf{B}\cdot d\mathbf{S}_k = \tfrac{1}{2}\sum_{k=1}^{n} J_k\Phi_k, \qquad (10\text{–}12)$$

where we have made use of Stokes' theorem, and $\Phi_k = \int \mathbf{B}\cdot d\mathbf{S}_k$ is the flux linking the kth circuit, as in Section 8–4 and Eq. (9–1). To derive the forces, let us assume that the ith circuit is subjected to a virtual, infinitesimally slow velocity $\mathbf{u}_i$. Then the rate at which a force $\mathbf{F}_i$ acting on the ith circuit is doing work is given by $\mathbf{F}_i\cdot\mathbf{u}_i$. The total rate of energy change in this virtual process is zero, but it is balanced between four quantities:

(1) Rate at which mechanical work is done on the ith circuit, $\mathbf{F}_i\cdot\mathbf{u}_i$.

(2) Rate of change of magnetic field energy, dU_m/dt.

(3) Rate of Joule heat losses, $\Sigma J_k^2 R_k$, where R_k is the resistance of the kth circuit.

(4) Rate at which work is being done on the electromotive forces within the circuits, $-\Sigma J_k\mathcal{E}_k$.

Thus we obtain

$$\mathbf{F}_i\cdot\mathbf{u}_i + \frac{dU_m}{dt} + \Sigma J_k^2 R_k - \Sigma J_k\mathcal{E}_k = 0. \qquad (10\text{–}13)$$

We are assuming that the magnetic field energy U_m is expressed explicitly as a function of the coordinates x_k of the kth current loop and of the current J_k as independent variables. Note that because of the Joule heat and battery work we cannot simply equate the force on the ith circuit, $\mathbf{F}_i$, to the negative gradient of the field energy U_m at constant current, a procedure that would be justified if no other energy terms were present.

Let us now consider a special type of constant current process, namely, that in which the external electromotive forces are adjusted as a function of the virtual velocity $\mathbf{u}_i$ corresponding to the change of a single parameter, x_i, so that the currents within the system remain constant. In this case, if we use Eq. (10–12) for U_m in Eq. (10–13) and write $\mathcal{E} = JR + d\Phi/dt$ according to Faraday's law [Eq. (9–2)], we obtain

$$\mathbf{F}_i\cdot\mathbf{u}_i + \frac{1}{2}\Sigma J_k\frac{d\Phi_k}{dt} - \Sigma J_k\frac{d\Phi_k}{dt} = 0$$

or

$$\mathbf{F}_i\cdot\mathbf{u}_i = \frac{\partial U_m}{\partial t} = \frac{\partial U_m}{\partial x_{i_\alpha}} u_{i_\alpha} \qquad (10\text{–}14)$$

under the conditions that all J's are constant and the positions of all circuits but the ith are held fixed. Hence,

$$F_{i_\alpha} = \left.\frac{\partial U_m}{\partial x_{i_\alpha}}\right|_{J \text{ constant}} \tag{10–15}$$

is the force exerted *by* the field *on* the ith conductor. Note that the sign is opposite to the sign which would be expected if other energy terms were neglected. This means that in order to maintain a constant current in the circuits, as the geometry changes, the batteries must do exactly twice as much work as that done against the external forces, in addition to supplying the Joule heat losses. Equation (10–15) is very useful for computing the forces acting on current-carrying circuits if the magnetic field energy is expressible in terms of the current producing the field.

10–3 Inductance. To express the magnetic field energy

$$U_m = \tfrac{1}{2}\sum_k J_k\Phi_k \tag{10–12}$$

as a function of current and geometry it is useful to introduce the concept of inductance. The flux through the kth circuit is given by

$$\Phi_k = \int \mathbf{B}\cdot d\mathbf{S}_k = \int (\nabla \times \mathbf{A})\cdot d\mathbf{S}_k = \oint_k \mathbf{A}\cdot d\mathbf{l}_k. \tag{10–16}$$

The general solution for **A**,

$$\mathbf{A} = \frac{\mu_0}{4\pi}\int \frac{\mathbf{j}}{r}\,dv \tag{8–29}$$

can be written as a line integral for linear circuits:

$$\mathbf{A}(x_k) = \frac{\mu_0}{4\pi}\sum_i \oint \frac{J_i\,d\mathbf{l}_i}{r_{ik}}. \tag{10–17}$$

With Eq. (10–17) substituted in Eq. (10–16), we may write

$$\Phi_k = \Sigma L_{ik}J_i, \tag{10–18}$$

where

$$L_{ik} = \frac{\mu_0}{4\pi}\oint\oint \frac{d\mathbf{l}_i\cdot d\mathbf{l}_k}{r_{ik}} = L_{ki} \tag{10–19}$$

is a purely geometrical quantity. L_{ik} is called the mutual inductance between the ith and kth circuits, and Eq. (10–19) is known as Neumann's formula.

Similarly, Eq. (10–12) becomes

$$U_m = \tfrac{1}{2}\sum_i\sum_k L_{ik}J_iJ_k. \tag{10–20}$$

The force acting on the ith circuit is thus, from Eq. (10-15),

$$F_{i_\alpha} = + \left.\frac{\partial U_m}{\partial x_{i_\alpha}}\right|_{J \text{ constant}} = \tfrac{1}{2} \sum_j \sum_k J_j J_k \frac{\partial L_{jk}}{\partial x_{i_\alpha}}. \tag{10-21}$$

In the sums of Eqs. (10-20) and (10-21) each term for which the two indices are different occurs twice, while that for equal indices occurs only once. The mutual energy of two circuits is therefore

$$U_m = J_1 J_2 L_{12}, \tag{10-22}$$

while the self-energy of the two circuits is

$$U_m = \tfrac{1}{2}(J_1^2 L_{11} + J_2^2 L_{22}), \tag{10-23}$$

where the L_{ii} are called the self-inductances.

The force expression (10-21) is in agreement with the original magnetic interaction of Eq. (7-21). If we substitute Eq. (10-19) in the force equation as applied to two circuits, we obtain

$$\begin{aligned} \mathbf{F}_1 &= \frac{\mu_0}{4\pi} J_1 J_2 \oint\oint (d\mathbf{l}_1 \cdot d\mathbf{l}_2)\, \nabla \frac{1}{r_{12}} \\ &= -\frac{\mu_0}{4\pi} J_1 J_2 \oint\oint \mathbf{r}_{12} \frac{d\mathbf{l}_1 \cdot d\mathbf{l}_2}{r_{12}^3}, \end{aligned}$$

which is identical with Eq. (7-24).

The force equation (10-21) may be written in the simple form

$$F_{i_\alpha} = J_i \Sigma_k J_k \frac{\partial L_{ik}}{\partial x_{i_\alpha}}, \tag{10-24}$$

since only the terms for which either $j = i$ or $k = i$ depend on x_i and give nonzero derivatives. Equation (10-24) can also be written as

$$F_{i_\alpha} = J_i \left.\frac{\partial \Phi_i}{\partial x_{i_\alpha}}\right|_{J \text{ constant}}, \tag{10-25}$$

where Φ_i is the flux linking the ith circuit. This expression is in evident agreement with the elementary force relation

$$d\mathbf{F} = J\, d\mathbf{l} \times \mathbf{B}. \tag{10-26}$$

These considerations enable us to express the general variation of the field energy as a function, independently, of the geometrical coordinates and of the currents. Since

$$U_m = \tfrac{1}{2} \sum_i \sum_k L_{ik} J_i J_k, \tag{10-20}$$

we have (note that each term occurs twice):

$$\delta U_m = \sum_i \left[\left(\sum_k L_{ik} J_k \right) \delta J_i + \sum_k \frac{\partial L_{ik}}{\partial x_{i_\alpha}} \delta x_{i_\alpha} J_i J_k \right]. \tag{10-27}$$

Hence, from Eqs. (10–18) and (10–24),

$$\delta U_m = \sum_i (\Phi_i \, \delta J_i + F_{i_\alpha} \, \delta x_{i_\alpha}), \tag{10-28}$$

where Φ_i is the total flux linking the ith circuit. It may be seen that positions and currents play the roles of extensive variables in the thermodynamic sense, while the forces and flux play the roles of intensive variables.

The "back emf" terms can be ignored in force calculations if the flux linkages are held constant; in that case the work supplied by the battery exactly balances the Joule heat loss. In these circumstances, directly from Eq. (10–13), we have

$$F_{i_\alpha} = - \left. \frac{\partial U_m}{\partial x_{i_\alpha}} \right|_{\Phi \text{ constant}}, \tag{10-29}$$

in contrast to Eq. (10–15).

The self-inductances and mutual inductances can be calculated by several means other than Neumann's formula. One method is to use the defining equation, Eq. (10–18). The flux linking the kth circuit due to the current in the ith circuit can be evaluated directly from the known field or vector potential of the ith circuit. A second method, which is particularly useful when continuous current distributions and therefore partial flux linkages are involved, makes explicit use of the two expressions for the magnetic field,

$$U_m = \tfrac{1}{2} \sum_{i=k}^{n} \sum_{k=1}^{n} L_{ik} J_i J_k = \tfrac{1}{2} \int \mathbf{H} \cdot \mathbf{B} \, dv. \tag{10-30}$$

Solutions for $\mathbf{H}$ and $\mathbf{B}$ may be obtained by methods already discussed. The calculation of inductances is then carried out by evaluating the integrals of $\mathbf{H} \cdot \mathbf{B}$.

In computing the inductances of current-carrying conductors it is usually advantageous to separate the problem into two parts: (1) the inductance associated with the field outside the conductor, and (2) the contribution to the inductance by the field energy inside the wire. It is necessary to make this separation because the inductance due to the external field cannot be computed under the assumption of zero radius for the conductor: such an assumption will generally lead to a logarithmic divergence of the integral involved. We shall see that at high frequencies the contribution to the inductance by the field within the wire must become negligible, since the currents do not penetrate into the wire. At lower frequencies, and particularly in case the conductors are ferromagnetic, this internal term may make an appreciable contribution.

10–4 Magnetic volume force. We have now calculated the forces between current systems in terms of the currents and the appropriate geometrical quantities. These forces are, as would be expected, re-expressions of the original Ampere interaction law given in Eq. (7–21). We can also, in analogy to the electrostatic case, derive an expression for the magnetic body force per unit volume in terms of the field, the permeability, and the current at a given point. In the electrostatic case we defined $\mathbf{F}_v$ by the relation

$$\frac{dU}{dt} = -\int \mathbf{F}_v \cdot \mathbf{u}\, dv \tag{6–23}$$

and found that $\mathbf{F}_v$ was given by

$$\mathbf{F}_v = \rho\mathbf{E} - \frac{\epsilon_0}{2} E^2 \nabla\kappa + \frac{\epsilon_0}{2} \nabla\left(E^2 \frac{d\kappa}{dg} g\right). \tag{6–40}$$

Similarly, we may write

$$\frac{dU_m}{dt} = -\int \mathbf{F}_v \cdot \mathbf{u}\, dv$$

and identify $\mathbf{F}_v$ as the magnetic body force. Instead of writing the general expression, let us make the following restrictions: (1) The medium is linear, i.e., its permeability is not a function of the field. (2) There is no permanent magnetic moment present. (3) There is no magnetostriction, i.e., $d\kappa_m/dg = 0$. Under these conditions a straightforward calculation yields

$$\mathbf{F}_v = \mathbf{j} \times \mathbf{B} - \frac{1}{2\mu_0} B^2 \nabla\kappa_m. \tag{10–31}$$

It is again possible, in accord with the requirements of a satisfactory field theory, to derive the total force on a given volume in terms of the value of the field on the boundary of this volume. In other words, it is possible to define a stress tensor from which the volume force is derivable by the tensor divergence relation of Eq. (6–41). The form of the Maxwell tensor in the magnetic case, in the absence of a magnetostriction term, is formally analogous to Eq. (6–54):

$$T_{\alpha\beta} = H_\alpha B_\beta - \frac{\delta_{\alpha\beta}}{2} H_\gamma B_\gamma, \tag{10–32}$$

with the customary summation convention. The geometrical interpretation of this tensor leads to the same conclusions about magnetic forces as were reached in Chapter 6 about electrical forces: the direction of the magnetic field bisects the angle between the normal to a surface and the direction of the resultant magnetic stress that acts on the surface. The magnitude of the magnetic stress normal to the field or parallel to the field is $HB/2$.

10–5 General expressions for electromagnetic energy. Thus far we have considered the electrostatic field and the magnetostatic field, or quasi-stationary current fields, separately. Let us now inquire as to which of the energy, force, or momentum relations will need modification for the general case in which no restrictions on the time rate of change of the field quantities are imposed. We can confine our attention to vacuum fields and nonpermeable conductors without omitting results of any great interest.

Maxwell's equations *in vacuo* were shown to be:

$$(1)\qquad \nabla \cdot \mathbf{D} = \rho_{\text{true}},$$

$$(2)\qquad \nabla \cdot \mathbf{B} = 0,$$

$$(3)\qquad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \qquad (9\text{–}7)$$

$$(4)\qquad \nabla \times \mathbf{H} = \mathbf{j}_{\text{true}} + \frac{\partial \mathbf{D}}{\partial t}.$$

These equations completely represent the behavior of electromagnetic fields when they are considered in combination with suitable constitutive equations and boundary conditions. This is true even for rapidly varying fields, or at least no internal contradiction is present if an arbitrary rate of change is assumed. Care must be taken in using constitutive equations, however, since the material constants are generally dependent on the frequency of the fields. We shall often restrict ourselves to vacuum conditions in order to avoid complications due to such special properties of the constitutive equations.

An energy integral of Maxwell's equations can be obtained as a result of taking the scalar product of the third and fourth of these equations with **H** and **E** respectively. If we subtract the two resulting equations and make use of the vector identity for the divergence of a cross product, Eq. (10–3), we obtain

$$\nabla \cdot (\mathbf{E} \times \mathbf{H}) = -\frac{\partial \mathbf{B}}{\partial t} \cdot \mathbf{H} - \mathbf{E} \cdot \mathbf{j} - \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t}. \qquad (10\text{–}33)$$

Let us now take the volume integral of Eq. (10–33), assuming that the constitutive equations are linear:

$$-\frac{\partial}{\partial t}\int \tfrac{1}{2}(\mathbf{H} \cdot \mathbf{B} + \mathbf{E} \cdot \mathbf{D})\, dv = \int \mathbf{E} \cdot \mathbf{j}\, dv + \int (\mathbf{E} \times \mathbf{H}) \cdot d\mathbf{S}, \qquad (10\text{–}34)$$

where the surface term arises as a result of using the divergence theorem. The left side of Eq. (10–34) is recognized as the rate of decrease of the sum of the electric and magnetic field energies [Eqs. (6–14) and (10–7)] as derived for the static case. By means of Eq. (10-1) the first term on the right can be written

$$\int \mathbf{E}\cdot\mathbf{j}\,dv = \int\left(\frac{j^2}{\sigma} - \mathbf{E}'\cdot\mathbf{j}\right)dv. \tag{10-35}$$

Equation (10-35) therefore represents the sum of the Joule heat losses and the negative rate at which the electromotive forces are doing work.

The last term on the right side of Eq. (10-34) demands careful consideration. It represents an energy which has been hitherto neglected, since for static and quasi-static fields the integral can be made to vanish by taking an arbitrarily large enclosing surface. As we shall see later, however, the electric and magnetic radiation fields of charge motions and currents fall off in general only as $1/r$ at large distances. In that case $\int(\mathbf{E}\times\mathbf{H})\cdot d\mathbf{S}$ will approach a constant value for an arbitrarily large surface, and thus the integral may contribute to the energy balance.

The vector

$$\mathbf{N} = \mathbf{E}\times\mathbf{H} \tag{10-36}$$

is known as the Poynting vector. In terms of Eq. (10-34) it can be considered to represent the electromagnetic field energy flow per unit area per unit time across a given surface. It must be noted, however, that only the entire surface integral of **N** contributes to the energy balance. Paradoxical results may be obtained if one tries to identify the Poynting vector with the energy flow per unit area at any particular point. The energy term arose as the volume integral of $\nabla\cdot(\mathbf{E}\times\mathbf{H})$, and the net energy flow in the electromagnetic field will always vanish if the divergence of the Poynting vector is zero. For example, in static superposed electric and magnetic fields we may have nonzero values of the Poynting vector at various points in space, but its divergence vanishes everywhere.

We have now seen that Eq. (10-34) can be taken to represent the overall energy balance between the electric and magnetic energy of the field, the loss due to resistive heating, the work done by sources of electromotive force, and the radiation loss. It appears that the expressions derived for the energy densities of the electrostatic and magnetostatic fields retain their validity when the fields are allowed to vary arbitrarily with the time. The only additional consideration needed in order to conserve energy is to assume that fields may carry energy in or out of the volume of integration at a rate that is given by the surface integral of the Poynting vector.

A new consequence of the introduction of the surface term into the conservation laws is the possibility of balancing energy and momentum over only part of a system. The surface integral over the Poynting vector permits us to conserve energy in one part of a system whether radiative processes are present or not. To illustrate this point, consider the simple process of a battery (electromotive field $\mathbf{E}'$) feeding a current (density $\mathbf{j}$) to a resistor (of conductivity σ), as in Fig. 10–1. Let us take the energy balance over the volume of length l of the resistor, whose radius a has been magnified in the drawing. By elementary considerations,

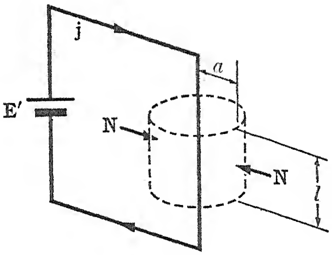


FIG. 10–1 Illustrating the role of the Poynting vector in the conservation of energy over a part of a circuit.

$$E = j/\sigma; \quad H = \tfrac{1}{2}ja; \quad N = \frac{j^2 a}{2\sigma}$$

where $\mathbf{N}$ is directed inward over the lateral surface of the cylinder. Hence

$$\int \mathbf{N} \cdot d\mathbf{S} = -j^2 \frac{(\pi a^2 l)}{\sigma} = -\int \frac{j^2}{\sigma}\, dv = -\text{ Joule heat.}$$

Equation (10–34) is thus satisfied, and energy is balanced by considering the *field* as "feeding" the resistor via the Poynting vector, without the explicit introduction of the source of energy, i.e., the battery field $\mathbf{E}'$.

10–6 Momentum balance. In order to investigate the momentum relations in the electromagnetic fields, we shall inquire as to whether the divergence of the complete Maxwell stress tensor will give a volume force which is in accord with experience. If we neglect the electrostriction and magnetostriction terms the stress tensor is the sum of Eqs. (6–53) and (10–32):

$$T_{\alpha\beta} = E_\alpha D_\beta - \tfrac{1}{2}\delta_{\alpha\beta}E_\gamma D_\gamma + H_\alpha B_\beta - \tfrac{1}{2}\delta_{\alpha\beta}H_\gamma B_\gamma. \qquad (10\text{–}37)$$

The tensor divergence of Eq. (10–37) is

$$\frac{\partial T_{\alpha\beta}}{\partial x_\beta} = E_\alpha \frac{\partial D_\beta}{\partial x_\beta} + D_\beta \frac{\partial E_\alpha}{\partial x_\beta} - \frac{1}{2} E^2 \epsilon_0 \frac{\partial \kappa}{\partial x_\alpha} - D_\beta \frac{\partial E_\beta}{\partial x_\alpha}$$
$$+ H_\alpha \frac{\partial B_\beta}{\partial x_\beta} + B_\beta \frac{\partial H_\alpha}{\partial x_\beta} - \frac{1}{2} H^2 \mu_0 \frac{\partial \kappa_m}{\partial x_\alpha} - B_\beta \frac{\partial H_\beta}{\partial x_\alpha}. \qquad (10\text{–}38)$$

Equation (10–38) can be expressed in vector form: by Maxwell's first equation $\nabla \cdot \mathbf{D} = \rho_{\text{true}}$ in the first term; the second and fourth terms become $-\mathbf{D} \times (\nabla \times \mathbf{E})$, which is equal to $\mathbf{D} \times \partial\mathbf{B}/\partial t$; the sixth and eighth terms become $-\mathbf{B} \times (\nabla \times \mathbf{H})$, which is equal to $-\mathbf{B} \times (\mathbf{j}_{\text{true}} + \partial\mathbf{D}/\partial t)$; the fifth term vanishes due to the fact that $\nabla \cdot \mathbf{B} = 0$; and the third and seventh terms involve merely gradients of κ and κ_m. Therefore

$$\frac{\partial T_{\alpha\beta}}{\partial x_\beta} = \left[\mathbf{E}\rho_{\text{true}} - \frac{\epsilon_0}{2} E^2 \nabla\kappa - \frac{\mu_0}{2} H^2 \nabla\kappa_m - \mathbf{B} \times \mathbf{j}_{\text{true}} + \frac{\partial}{\partial t}(\mathbf{D} \times \mathbf{B})\right]_\alpha . \qquad (10\text{–}39)$$

This entire tensor divergence may be expressed as the sum of two terms:

$$\frac{\partial T_{\alpha\beta}}{\partial x_\beta} = \left[\mathbf{F}_{ev} + \frac{\partial}{\partial t}(\mathbf{D} \times \mathbf{B})\right]_\alpha, \tag{10-40}$$

where the first term,

$$\mathbf{F}_{ev} = \mathbf{E}\rho_{\text{true}} - \frac{\epsilon_0}{2} E^2 \nabla \kappa - \frac{\mu_0}{2} H^2 \nabla \kappa_m - \mathbf{B} \times \mathbf{j}_{\text{true}} \tag{10-41}$$

is the ordinary volume force acting on material bodies in a quasi-stationary electromagnetic field. Equation (10-41) accounts fully for the volume force on true charges or inhomogeneous dielectrics in an electric field, and for that on true currents or inhomogeneous permeable material in a magnetic field. The second term in Eq. (10-40) is new, and is proportional to the time rate of change of the Poynting vector. Equation (10-40) may be written explicitly

$$\frac{\partial T_{\alpha\beta}}{\partial x_\beta} = \left[\mathbf{F}_{ev} + \mu\epsilon \frac{\partial(\mathbf{E} \times \mathbf{H})}{\partial t}\right]_\alpha$$

$$= \left[\mathbf{F}_{ev} + \frac{\kappa\kappa_m}{c^2}\frac{\partial \mathbf{N}}{\partial t}\right]_\alpha, \tag{10-42}$$

where

$$\epsilon_0\mu_0 = 1/c^2. \tag{10-43}$$

The existence of $\mathbf{F}_{ev}$ depends on the presence of material bodies carrying charges or endowed with dielectric properties. On the other hand, the second term in Eq. (10-42) does not vanish even *in vacuo*, and therefore it would superficially suggest the idea of a volume force on the vacuum. This term has evoked a great deal of speculation. It fits into an ether theory in which vacuum possesses various mechanical properties that enable it to transmit elastic waves and also to sustain body forces. The only way such an ether force could be measured would be by means of the action of the ether on matter.

According to Lorentz' electron theory, however, the only force which has physical significance is a resultant force which arises from the space-time average forces acting on material charges and currents, namely, those obtained by averaging

$$\mathbf{F} = \rho(\mathbf{E} + \mathbf{u} \times \mathbf{B}). \tag{10-44}$$

We shall also find that within the framework of the special theory of relativity no measurement can be devised which can determine the velocity or other properties of the ether. If therefore we adopt the point of view that the only volume force which has a place in physical theory is a force derivable from the Lorentz force, Eq. (10–44), it follows that the second term in Eq. (10–42) must be subtracted out. We then have for the volume force, when $\kappa = \kappa_m = 1$,

$$F_{v\alpha} = \frac{\partial T_{\alpha\beta}}{\partial x_\beta} - \frac{1}{c^2}\frac{\partial N_\alpha}{\partial t}, \tag{10–45}$$

which is equal to the Lorentz force.

If we apply this equation to a volume containing both matter and radiation, and bounded by a finite surface, we obtain*

$$F_\alpha = \int T_{\alpha\beta}\,dS_\beta - \frac{1}{c^2}\frac{\partial}{\partial t}\int N_\alpha\,dv. \tag{10–46}$$

Since the integrated body force F_α represents the total rate of change of mechanical momentum, p_α, of the volume, Eq. (10–46) can be written

$$\frac{d}{dt}\left[p_\alpha + \frac{1}{c^2}\int N_\alpha\,dv\right] = \int T_{\alpha\beta}\,dS_\beta. \tag{10–47}$$

This equation states that the sum of the rate of change of the mechanical momentum, plus a term equal to the volume integral of the Poynting vector divided by c^2, is equal to the surface integral of the total Maxwell stress transmitted across the surface surrounding the volume. If it were possible to choose a surface so large that it is in field-free space, then the sum of the mechanical momentum and the volume integral of the Poynting vector would be constant in time. This implies that the correction term, whose introduction into Eq. (10–45) was required only by our demand for a physical interpretation of the volume force, makes necessary a change in our concept of momentum.

In the absence of measurable physical properties for the ether, we are thus forced to modify the law of the conservation of momentum. It must apply not just to matter alone, but must also include a momentum density of the electromagnetic radiation field which is equal to the Poynting vector divided by a constant, c^2. The Poynting vector therefore appears in a dual role, as carrying energy and also as carrying momentum. It will turn out in the special theory of relativity that the transfer of energy corresponds to a transfer of momentum in the proportions that have been derived here. Actually, this is a property of all forms of energy flow, and is not confined to electromagnetic radiation.

---

* We have omitted a term which, in matter, is given by $\frac{(\kappa\kappa_m - 1)}{c^2}\frac{\partial}{\partial t}\int N_\alpha\,dv$, and is actually present when an electromagnetic wave travels through matter. Its net impulse due to a finite wave train always vanishes.

SUGGESTED REFERENCES

M. Abraham and R. Becker, *Classical Theory of Electricity and Magnetism.* In Chapters VIII and IX the energy theorems are developed in a clear and elementary way.

W. R. Smythe, *Static and Dynamic Electricity.* Characteristically rich in detailed applications to specific problems.

J. H. Jeans, *Electricity and Magnetism.* Again the organization of subject matter is very different from ours, but there are interesting applications.

EXERCISES

1. Use Neumann's formula to compute the mutual inductance per unit length between two equal parallel conductors. If the two wires constitute a single circuit, find the force between them.

2. Consider a coaxial cable consisting of a center wire of radius a and a thin sheath of radius $b \gg a$. Find the self-inductance per unit length of the pair of conductors as a circuit, and the mutual inductance of the core and sheath. Show that the mutual inductance of the core and sheath is the same as the self-inductance of the sheath. Are the answers altered if the core and sheath are not coaxial?

3. Find the torque on a solid conducting cylinder rotating slowly in a uniform magnetic field perpendicular to the axis of the cylinder.

4. A thin spherical shell of conductivity σ, thickness t, radius a, rotates with uniform angular velocity about an axis perpendicular to a uniform magnetic field. Calculate the power needed to maintain the rotation.

5. A steel sphere of radius a is magnetized permanently and uniformly with magnetization $\mathbf{M}$. A circular coil of N turns and radius $b > a$, carrying a current J, is mounted in a plane parallel to $\mathbf{M}$ so that the center of the coil coincides with the center of the sphere. Find the torque on the coil.

6. Find the total field energy, and the field energy density in the iron and air gap of the incomplete iron ring of Exercise 10, Chapter 8. Discuss the energy balance of the process of widening the gap slightly by means of an external mechanical force.

CHAPTER 11

THE WAVE EQUATION AND PLANE WAVES

11–1 The wave equation. In the energy integral of Maxwell's equations, Eq. (10–34), we interpreted a nonvanishing integral over a distant surface as representing a flow of energy which we called radiation. The equations may be combined to exhibit these propagating fields more explicitly: the four first order linear partial differential field equations can be reduced to a system of two second order linear partial differential equations. To make this reduction, let us recall that Maxwell's equations are

$$\begin{aligned} &(1) \qquad & \nabla\cdot\mathbf{D} &= \rho_{\text{true}}, \\ &(2) & \nabla\cdot\mathbf{B} &= 0, \\ &(3) & \nabla\times\mathbf{E} &= -\frac{\partial\mathbf{B}}{\partial t}, \\ &(4) & \nabla\times\mathbf{H} &= \mathbf{j}_{\text{true}}+\frac{\partial\mathbf{D}}{dt}. \end{aligned} \tag{9–7}$$

Consider a region where there are no true charges and no sources of electromotive force, so that $\rho_{\text{true}} = 0$ and $\mathbf{E}' = 0$, and where ϵ and μ are not functions of the coordinates or of the time. Take the curl of Eq. (9–7) (3) and substitute $\mu\mathbf{H}$ for $\mathbf{B}$:

$$\nabla\times(\nabla\times\mathbf{E}) = -\frac{\partial}{\partial t}(\nabla\times\mu\mathbf{H}). \tag{11–1}$$

On substituting Eq. (9–7) (4) for curl $\mathbf{H}$, we obtain

$$\nabla\times(\nabla\times\mathbf{E}) = -\mu\frac{\partial}{\partial t}\left(\mathbf{j}_{\text{true}}+\epsilon\frac{\partial\mathbf{E}}{\partial t}\right). \tag{11–2}$$

Now for any vector it is true that

$$\nabla\times(\nabla\times\mathbf{E}) = \nabla(\nabla\cdot\mathbf{E}) - \nabla^2\mathbf{E},$$

but $\nabla\cdot\mathbf{E} = 0$ in the charge-free region, and by Eq. (7–3) $\mathbf{j} = \sigma\mathbf{E}$. Therefore Eq. (11–2) becomes

$$\nabla^2\mathbf{E} - \frac{\kappa\kappa_m}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2} - \mu\sigma\frac{\partial\mathbf{E}}{\partial t} = 0, \tag{11–3}$$

where we have written $1/c^2$ for the product of free space constants $\mu_0\epsilon_0$. It can be seen from Eq. (11–3) that c must have the dimensions of velocity.

Equation (11–3) is known as the general wave equation. Usually either the second or the third term drops out in any particular application of this equation. In a nonconducting medium the third term vanishes, leaving a propagation equation for waves that travel with the velocity $u = 1/\sqrt{\mu\epsilon}$, as we shall see. In a conducting medium the second term is usually negligible, and we are left with the same differential equation as for heat conduction or diffusion. The relative magnitude of the two terms can be easily estimated for a field that varies with an angular frequency ω:

$$\mathbf{E} = \mathbf{E}(x_\alpha)e^{-i\omega t}. \tag{11–4}$$

When Eq. (11–4) is substituted in Eq. (11–3), we obtain an equation that does not depend on the time:

$$\nabla^2\mathbf{E} + \frac{\kappa\kappa_m}{c^2}\,\omega^2\mathbf{E} + i\mu\sigma\omega\mathbf{E} = 0. \tag{11–5}$$

Equation (11–5) may be written

$$\nabla^2\mathbf{E} + \left(1 + \frac{i\sigma}{\epsilon\omega}\right)\mu\epsilon\omega^2\mathbf{E} = 0, \tag{11–6}$$

and the relaxation time of the medium, discussed in Section 7–4, is given by

$$\tau = \frac{\epsilon}{\sigma}. \tag{7–20}$$

Hence we see that if the relaxation time τ is long compared with the period $2\pi/\omega$ of the sinusoidal vibration, so that the imaginary term in Eq. (11–6) can be neglected, we are left with a propagation equation. On the other hand, if the relaxation time is short compared with the period, then the imaginary term is large compared with unity and we have essentially a diffusion condition.

For all pure metals the relaxation time is of the order of 10^{-14} sec, and hence the diffusion type of equation is valid for all frequencies lower than those of the optical spectrum. This means that within metallic conductors the propagation term $\mu\epsilon\,\partial^2\mathbf{E}/\partial t^2$ can be neglected even in the ultra-high frequency radio region. To put it in a different way, the displacement current is negligible relative to the conduction current in metals at the highest frequencies that are theoretically attainable with macroscopic oscillators.

The coefficients that appear in Eq. (11–5) can be expressed in a variety of useful ways. If, for example, we write

$$\frac{\omega}{c} = \frac{1}{\lambda\!\!\!^{-}_0}, \tag{11–7}$$

$\lambda\!\!\!^{-}_0$ is the free space wavelength divided by 2π. Similarly, for a medium characterized by κ and κ_m, we may write

$$\frac{\omega}{u} = \frac{1}{\lambda\!\!\!^{-}}, \tag{11-8}$$

where $u = c/\sqrt{\kappa\kappa_m}$. The ratio of the magnitude of the conduction current to the magnitude of the displacement current can be rewritten in terms of these relations as

$$\frac{1}{\omega\tau} = \frac{\sigma}{\omega\epsilon} = \frac{\sigma}{\epsilon}\frac{\lambda\!\!\!^{-}}{u} = \sigma\lambda\!\!\!^{-}\sqrt{\frac{\mu}{\epsilon}} = \sigma\lambda\!\!\!^{-} R_0\sqrt{\frac{\kappa_m}{\kappa}}, \tag{11-9}$$

where

$$R_0 = \sqrt{\frac{\mu_0}{\epsilon_0}} \tag{11-10}$$

is a resistance whose numerical value is 376.7 ohms. This number is sometimes called the characteristic impedance of free space. The analogy between wave propagation and "lumped" transmission line parameters is useful in some practical considerations, but we shall here avoid such analogies in favor of considering the characteristics of the electromagnetic fields themselves. The significance of Eq. (11–9) might be stated by saying that if the resistance of a cube whose edge is $\lambda\!\!\!^{-}$ is larger than R_0, then in such a medium the displacement current is dominant, while if the reverse is true the conduction current governs the behavior of electromagnetic fields in the medium.

Equation (11–3) and the analogous equation for **H** or **B** are homogeneous in the field vectors because they were derived subject to the condition of no true charges and no sources of electromotive force. To investigate the relation between these fields and their sources it is convenient, as in the static case, to introduce potentials as intermediary field quantities. We shall return to this problem in Chapter 13. A number of interesting and important consequences that are independent of the origin of the fields can be considered without reference to the potentials.

11–2 Plane waves. Let us consider the case in which all fields are functions only of the distance of a given plane from the origin, as in Fig. 11–1. If this distance is ζ and if **n** is a unit vector normal to the plane, then for all points on the plane the spatial derivatives are functions of ζ only and

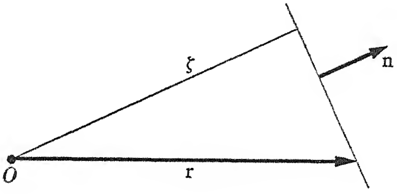


FIG. 11–1 All fields are constant on such plane surfaces as that indicated by the unit normal **n**.

the operator ∇ becomes

$$\nabla = \mathbf{n}\frac{\partial}{\partial \zeta}.$$

Thus Maxwell's equations reduce to

$$\begin{aligned}
&(1) && \mathbf{n}\cdot\frac{\partial \mathbf{D}}{\partial \zeta} = 0,\\
&(2) && \mathbf{n}\cdot\frac{\partial \mathbf{B}}{\partial \zeta} = 0,\\
&(3) && \mathbf{n}\times\frac{\partial \mathbf{E}}{\partial \zeta} = -\frac{\partial \mathbf{B}}{\partial t},\\
&(4) && \mathbf{n}\times\frac{\partial \mathbf{H}}{\partial \zeta} = j + \frac{\partial \mathbf{D}}{\partial t} = \sigma\mathbf{E} + \frac{\partial \mathbf{D}}{\partial t}.
\end{aligned} \tag{11-11}$$

By taking the scalar product of **n** and Eq. (11–11) (4), we see that

$$\mathbf{n}\cdot\left(\frac{\sigma}{\epsilon} + \frac{\partial}{\partial t}\right)\mathbf{D} = 0. \tag{11-12}$$

Equations (11–12) and (11–11) (1) imply that the longitudinal components of **D** and **E**, i.e., those components that are perpendicular to the plane surface of the figure, are independent of ζ and that their time dependence follows an exponential decay law in accordance with the characteristic relaxation time of the medium. Thus

$$E_n = E_{n_0}e^{-t/\tau} = E_{n_0}e^{-\frac{\sigma t}{\epsilon}}. \tag{11-13}$$

This means that the only longitudinal solution of the field equations is an electrostatic solution, and that in the presence of finite conductivity the electrostatic solutions will vanish exponentially with time.

Scalar multiplication of **n** and Eq. (11–11) (3) leads to

$$\mathbf{n}\cdot\frac{\partial \mathbf{B}}{\partial t} = 0, \tag{11-14}$$

which, together with Eq. (11–11) (2), shows that the only longitudinal component of the magnetic field compatible with the field equations is a stationary uniform magnetic field. Thus if there is a nonstatic solution it must be composed of transverse fields, fields whose vectors lie parallel to the plane of Fig. 11–1.

Each transverse component of the electric field obeys the differential equation

$$\frac{\partial^2 \mathbf{E}}{\partial \zeta^2} - \mu\epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} - \mu\sigma \frac{\partial \mathbf{E}}{\partial t} = 0, \tag{11-15}$$

which may be derived by eliminating **H** between Eqs. (11-11) (3) and (4) or may be written down directly from Eq. (11-3). Equation (11-15) is called the "equation of telegraphy." If the medium has zero conductivity, this equation reduces to

$$\frac{\partial^2 \mathbf{E}}{\partial \zeta^2} - \mu\epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0. \tag{11-16}$$

The general solution of Eq. (11-16) is

$$E = f(\zeta - ut) + g(\zeta + ut), \tag{11-17}$$

where f and g are arbitrary functions of their respective arguments, and $u = c/\sqrt{\kappa\kappa_m} = 1/\sqrt{\mu\epsilon}$, as before. The general solution represents waves traveling along ζ, i.e., a disturbance of form f is propagated in the positive ζ-direction as time progresses, and g is propagated in the opposite direction. u is the velocity with which the form of the disturbance travels, and is called the phase or wave velocity. The ratio of the free-space velocity to that in a medium characterized by κ and κ_m, $c/u = \sqrt{\kappa\kappa_m}$, is called the index of refraction of the medium, and is often designated by n.

If **E** is assumed to have a sinusoidal time variation, given by the factor $e^{-i\omega t}$, with an angular frequency ω, the solution of Eq. (11-16) is

$$\mathbf{E} = \mathbf{E}_0 e^{-i(\omega t \pm k\zeta)}, \tag{11-18}$$

where

$$k = \omega/u = 2\pi/\lambda = 1/\lambda\!\!\!^{-}.$$

It is often useful to write k as a vector in the direction of propagation. The propagation vector is given by

$$\mathbf{k} = \frac{\omega}{u^2}\mathbf{u} = \frac{1}{\lambda\!\!\!^{-}}\frac{\mathbf{u}}{u}, \tag{11-19}$$

so that a solution of Eq. (11-16) may be written as

$$\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)}. \tag{11-20}$$

Let us consider the case of the negative sign in the complex exponential of Eq. (11-18), so as to have a wave moving toward greater positive ζ. Equation (11-11) (3) becomes

$$\mathbf{n} \times \frac{\partial \mathbf{E}}{\partial \zeta} = i\omega\mathbf{B} = ik\mathbf{n} \times \mathbf{E}$$

or

$$\mathbf{B} = \sqrt{\mu\epsilon}\,\mathbf{n} \times \mathbf{E} = \frac{\mathbf{k}}{\omega} \times \mathbf{E}. \tag{11-21}$$

Thus we see that associated with each transverse component of **E** there is a magnetic field. If the direction of **E** or **B** is constant in time the wave is said to be plane polarized, or linearly polarized. Since the field equations show no correlation between different components of any one field vector, we lose no generality in restricting our discussion to a plane polarized wave. **E** and **B** are perpendicular to each other as well as to the direction of propagation, and the equations indicate that **E**, **B**, and **n** (or **k**) denote a right-handed coordinate system, in that order.

It is clear that the complex exponential time dependence for the fields leads to a particularly simple form for Maxwell's equations, and complex expressions such as that for **E** in Eq. (11–20) are, in general, very convenient mathematically, but we must remember that the actual fields are the real part of the complex quantity. The energy density associated with sinusoidally varying fields pulsates in the time, since it involves the factor $E^2 = E_0^2 \cos^2 (k\zeta - \omega t)$, the square of the real part of Eq. (11–20). The time average of E^2 is just $\frac{1}{2}E_0^2$. In general, the time average of the product of the real parts of two vectors both of which vary as $e^{-i\omega t}$ is one-half the real part of the product of one vector and the complex conjugate of the other. Thus

$$\overline{(\mathrm{Re}\,\mathbf{F}) \cdot (\mathrm{Re}\,\mathbf{G})} = \tfrac{1}{2}\,\mathrm{Re}\,(\mathbf{F} \cdot \mathbf{G}^*) = \tfrac{1}{2}\,\mathrm{Re}\,(\mathbf{F}^* \cdot \mathbf{G}), \tag{11-22}$$

where * denotes complex conjugate. The products on the right are independent of the time, so that the line denoting the average would have no significance. The verification of Eq. (11–22) follows readily from writing **F** and **G** as a sum of real and imaginary parts, and the details are left to a problem.

For the case of the plane wave fields given by Eqs. (11–20) and (11–21), and the related fields **D** and **H**, the time average of the energy density is

$$\overline{U} = \tfrac{1}{2}\overline{(\mathbf{H} \cdot \mathbf{B} + \mathbf{E} \cdot \mathbf{D})} = \tfrac{1}{2}\epsilon E_0^2 \text{ joules/meter}^3, \tag{11-23}$$

to which the electric and magnetic fields contribute equally. The Poynting vector, $\mathbf{N} = \mathbf{E} \times \mathbf{H}$, is in the direction of propagation of the wave, and its time average is

$$\begin{aligned}\overline{\mathbf{E} \times \mathbf{H}} &= \tfrac{1}{2}\sqrt{\epsilon/\mu}\, E_0^2\, \mathbf{n} \text{ watts/meter}^2 \\ &= \overline{U}\mathbf{u}.\end{aligned} \tag{11-24}$$

Thus the energy density associated with a plane wave in a stationary homogeneous nonconducting medium is propagated with the same velocity as that of the fields.

11–3 Radiation pressure. We have seen in Section 10–6 that the ether theory and the Lorentz electron theory lead to different expressions for the volume force. The divergence of the Maxwell stress tensor yields a volume force which does not vanish in the absence of charges and currents, and the correction needed to make this force consistent with the Lorentz force implies the existence of a momentum density in the electromagnetic field. Now we see that a plane wave carries energy. Let us next investigate the consequences of momentum balance applied to a plane wave.

Let us first consider a linearly polarized plane wave incident normally on a slab of material that absorbs it completely—a blackbody. If we take the x-axis as the direction of propagation, we may let y be the direction of the electric field and the magnetic field will be parallel to the z-axis. The Maxwell stress tensor,

$$T_{\alpha\beta} = E_\alpha D_\beta + H_\alpha B_\beta - \frac{\delta_{\alpha\beta}}{2}(E_\gamma D_\gamma + H_\gamma B_\gamma), \tag{10–37}$$

has in this case only the diagonal components

$$\begin{aligned} T_{xx} &= -\tfrac{1}{2}(E_y D_y + H_z B_z), \\ T_{yy} &= E_y D_y - \tfrac{1}{2}(E_y D_y + H_z B_z), \\ T_{zz} &= H_z B_z - \tfrac{1}{2}(E_y D_y + H_z B_z). \end{aligned} \tag{11–25}$$

The force derived by Eq. (10–45), which makes the volume force agree with the Lorentz force in a material medium, is

$$F_{v_\alpha} = \frac{\partial T_{\alpha\beta}}{\partial x_\beta} - \mu\epsilon \frac{\partial}{\partial t}(\mathbf{E} \times \mathbf{H})_\alpha, \tag{11–26}$$

of which the x-component is

$$-\frac{1}{2}\frac{\partial}{\partial x}(E_y D_y + H_z B_z) - \mu\epsilon \frac{\partial}{\partial t}(E_y H_z). \tag{11–27}$$

The y- and z-components of the force vanish, since the fields are not functions of y and z, and since the Poynting vector has no y- and z-components.

Since the fields are absorbed, the x-component of the time average volume force can be integrated to give the total pressure on the slab of material:

$$\int_0^\infty \overline{F_{v_x}}\, dx = \tfrac{1}{2}(\epsilon \overline{E^2} + \mu \overline{H^2}), \tag{11–28}$$

which is exactly equal to the energy density of the incoming radiation field. The time average of the time derivative of the Poynting vector vanishes if E_y and H_z vary sinusoidally, and therefore this term has been omitted from Eq. (11–28). If we wish to consider the net impulse transmitted by a wave train of finite length the pressure must be integrated from time

$-\infty$ to time $+\infty$, and the time derivative will also integrate out in this case; it can contribute only to the instantaneous value of the pressure during the absorption of a wave train, and not to the over-all effect. The radiation pressure is thus the same as in the ether theory—it is unaffected (except for transients) by the correction term needed to bring the volume force into accord with the Lorentz force. Note that Eq. (11–28) agrees with the pressure calculated by assuming that the momentum of volume $c\,dt$ given by $\mathbf{N}c\,dt/c^2$ is absorbed per unit area of the body.

This result can be summarized by saying that the phenomenon of radiation pressure is consistent with the concept of momentum of electromagnetic waves, and with the more general concept of momentum carried by any energy transmitting process. This pressure was predicted by the classical prerelativity ether theory of electromagnetic radiation, however, and is not changed by modern refinements of electromagnetic theory. The force involved is simply the Lorentz force, and all detailed models of the absorption process lead to the same answer, namely, that the pressure is equal to the energy density of the incident radiation. Measurements of radiation pressure have verified the theoretical predictions of its intensity.

Radiation pressure can also be computed for the case of wholly or partly reflecting surfaces and oblique incidence. Let us take the case of isotropic homogeneous radiation in a cavity, i.e., the radiation is unpolarized and of equal intensity in all directions. By symmetry it is evident that the only nonvanishing component of the volume force, Eq. (11–26), is that normal to the wall. If we take the x-axis normal to the element of surface under consideration, this force may be written out:

$$F_{v_x} = \epsilon\left[\frac{\partial E_x^2}{\partial x} + \frac{\partial}{\partial y}(E_xE_y) + \frac{\partial}{\partial z}(E_xE_z)\right] + \mu\left[\frac{\partial H_x^2}{\partial x} + \frac{\partial}{\partial y}(H_xH_y) + \frac{\partial}{\partial z}(H_xH_z)\right] - \frac{\epsilon}{2}\frac{\partial E^2}{\partial x} - \frac{\mu}{2}\frac{\partial H^2}{\partial x}. \tag{11–29}$$

(The time derivative term is omitted for the same reasons as before.) Due to the fact that different components of the electric field are uncorrelated, the time averages of all the cross terms vanish, and we obtain

$$\overline{F_{v_x}} = \epsilon\frac{\partial \overline{E_x^2}}{\partial x} - \frac{\epsilon}{2}\frac{\partial \overline{E^2}}{\partial x} + \mu\frac{\partial \overline{H_x^2}}{\partial x} - \frac{\mu}{2}\frac{\partial \overline{H^2}}{\partial x}. \tag{11–30}$$

Since, according to our assumption, the fields are oriented completely at random,

$$\overline{E_x^2} = \overline{E_y^2} = \overline{E_z^2} = \tfrac{1}{3}\overline{E^2}$$

and, similarly,

$$\overline{H_x^2} = \tfrac{1}{3}\overline{H^2}.$$

Therefore the normal force becomes

$$\overline{F_{v_x}} = -\frac{1}{6}\frac{\partial}{\partial x}[\epsilon\overline{E^2} + \mu\overline{H^2}] = -\frac{\partial}{\partial x}\left(\frac{\overline{U}}{3}\right),$$

where U is the energy density of the radiation. Integrating from the surface of the medium into field-free space, we obtain the result that the total radiation pressure is equal to $\frac{1}{3}$ of the energy density of the radiation:

$$\int_0^\infty \overline{F_{v_x}}\, dx = \tfrac{1}{3}\overline{U}. \tag{11-31}$$

This equation forms the basis for thermodynamic derivations of the Stefan-Boltzmann law and the Wien displacement law for blackbody radiation.

It should be noted that for the existence of completely homogeneous isotropic radiation, and thus of the randomness utilized in simplifying Eq. (11-29), it is essential that the walls be "black," i.e., perfectly absorbing. Equation (11-31) must not be assumed valid for other cases, as, for example, cavities with specularly reflecting walls.

11-4 Plane waves in a moving medium. Unlike radiation pressure, which was first predicted by electromagnetic theory, the velocity of light in a moving medium was derived by Fresnel on the assumption of elastic vibrations in a stationary ether, and confirmed by Fizeau as early as 1853. Fresnel's prediction, which will later be derived relativistically, is also in agreement with Maxwell's equations in moving media, provided that we interpret the velocity of the medium as that relative to the frame of reference in which the free-space velocity of light would be c.

According to Fresnel and Fizeau, the velocity of light in a body moving with a velocity $\mathbf{v}$ relative to the observer is given by

$$u = u_0 + \left(1 - \frac{1}{n^2}\right)\mathbf{v}\cdot\mathbf{n} \tag{11-32}$$

where $n = \sqrt{\kappa\kappa_m}$ is the index of refraction of the medium and $u_0 = c/n$. It should not be confused with the unit vector $\mathbf{n}$, which is in the direction of wave propagation, i.e., $\mathbf{u} = u\mathbf{n}$. We may derive the velocity of electromagnetic waves in a nonpermeable medium moving with a velocity small compared with c from the considerations of Section 9-4. Maxwell's equations in a moving medium free from true charges and currents become

$$\begin{aligned} &(1) \qquad \nabla\cdot\mathbf{D} = 0,\\ &(2) \qquad \nabla\cdot\mathbf{B} = 0, \end{aligned} \tag{11-33}$$

$$(3)\qquad \nabla\times\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t},$$

$$(4)\qquad \nabla\times\mathbf{B}/\mu_0 = \frac{\partial\mathbf{P}}{\partial t} + \epsilon_0\frac{\partial\mathbf{E}}{\partial t} + \nabla\times(\mathbf{P}\times\mathbf{v}), \qquad (11\text{–}33)$$

as may be seen from Eqs. (9–18). The polarization is produced by the effective field in the moving medium:

$$\mathbf{P} = \epsilon_0(\kappa-1)(\mathbf{E}+\mathbf{v}\times\mathbf{B}). \qquad (9\text{–}19)$$

Note that all the fields in Eqs. (11–33) and (9–19) are those which would be measured in the stationary frame of the observer. If we substitute the polarization field, Eq. (9–19), into Eq. (11–33) (4), and then make use of the third Maxwell equation, we obtain, correct to terms linear in $\mathbf{v}$:

$$\begin{aligned}\nabla\times\mathbf{B}&\\ &= \mu_0\left\{\epsilon_0\frac{\partial\mathbf{E}}{\partial t} + \epsilon_0(\kappa-1)\frac{\partial\mathbf{E}}{\partial t} + \epsilon_0(\kappa-1)\left[\mathbf{v}\times\frac{\partial\mathbf{B}}{\partial t} + \nabla\times(\mathbf{E}\times\mathbf{v})\right]\right\}\\ &= \frac{\kappa}{c^2}\left\{\frac{\partial\mathbf{E}}{\partial t} + \left(1-\frac{1}{\kappa}\right)[-\mathbf{v}\times(\nabla\times\mathbf{E}) + \nabla\times(\mathbf{E}\times\mathbf{v})]\right\}. \qquad (11\text{–}34)\end{aligned}$$

If we expand the vector triple products, noting that $\nabla\cdot\mathbf{E} = 0$, Eq. (11–34) becomes, for constant $\mathbf{v}$,

$$\nabla\times\mathbf{B} = \frac{\kappa}{c^2}\left\{\frac{\partial\mathbf{E}}{\partial t} + \left(1-\frac{1}{\kappa}\right)[2(\mathbf{v}\cdot\nabla)\mathbf{E} - \nabla(\mathbf{v}\cdot\mathbf{E})]\right\}. \qquad (11\text{–}35)$$

On taking the curl and making use of Eqs. (11–33)(3) and (2) we obtain

$$\nabla^2\mathbf{B} = \frac{\kappa}{c^2}\left[\frac{\partial^2\mathbf{B}}{\partial t^2} + 2\left(1-\frac{1}{\kappa}\right)(\mathbf{v}\cdot\nabla)\frac{\partial\mathbf{B}}{\partial t}\right] \qquad (11\text{–}36)$$

for the wave equation in slowly moving media.

Since we are here interested only in terms linear in the velocity, it is permissible to make the approximation that for plane wave propagation in the direction $\mathbf{n}$ along the coordinate ζ,

$$\mathbf{v}\cdot\nabla = (\mathbf{v}\cdot\mathbf{n})\frac{\partial}{\partial\zeta} = -\frac{\mathbf{v}\cdot\mathbf{n}}{u_0}\frac{\partial}{\partial t}.$$

Hence,

$$\nabla^2\mathbf{B} + \frac{1}{u_0^2}\frac{\partial^2\mathbf{B}}{\partial t^2}\left[1 - 2\left(1 - \frac{1}{\kappa}\right)\frac{\mathbf{v}\cdot\mathbf{n}}{u_0}\right] = 0. \tag{11-37}$$

This is a wave equation that corresponds to the propagation velocity

$$u = u_0\left[1 - \left(1 - \frac{1}{\kappa}\right)2\frac{\mathbf{v}\cdot\mathbf{n}}{u_0}\right]^{-1/2} \simeq u_0 + \left(1 - \frac{1}{\kappa}\right)\mathbf{v}\cdot\mathbf{n}, \tag{11-38}$$

which agrees to the order v/c with Eq. (11–32) and has been checked not only by Fizeau but also, with greater accuracy, by Michelson and Morley.

The physical interpretation of Eq. (11–38) is that the only part of the propagation velocity of the wave which is affected by the motion of the medium is that proportional to

$$1 - \frac{1}{\kappa} = \frac{\kappa - 1}{\kappa} \propto \frac{\partial\mathbf{P}/\partial t}{\partial\mathbf{D}/\partial t}. \tag{11-39}$$

This fraction is thus proportional to the ratio of the polarization current to the displacement current. Since a polarization current does actually correspond to the motion of dipoles, it is quite reasonable to assume that the portion of the wave corresponding to these dipoles will be affected by the velocity of the medium. We shall see that the effect of a medium on a plane wave is simply that the medium is polarized by the incident wave and that the resulting dipoles produce a wave which combines with the incident radiation in such a way as to correspond to the over-all phase velocity. It is this coherent retarded component which is being radiated from a moving source in this case, and which gives rise to the Fresnel-Fizeau coefficient. It may be remarked that the physical assumptions involved in this classical nonrelativistic theory are consistent with those of relativity, and we shall see that all these results are in accord with relativistic principles.

11–5 Reflection and refraction at a plane boundary. The boundary conditions inherent in the field equations determine what happens when a plane wave is incident on a boundary between media of different electric and magnetic properties. Let us consider a plane wave traveling in the direction of **k** in medium 1, and incident on the plane boundary between media 1 and 2, as shown in Fig. 11–2. The media are characterized by constants μ_1 and μ_2, ϵ_1 and ϵ_2. The propagation vector **k** and the unit vector **n** specifying the plane boundary determine what is called the plane of incidence of the wave. In accord with Eqs. (11–20) and (11–21) the

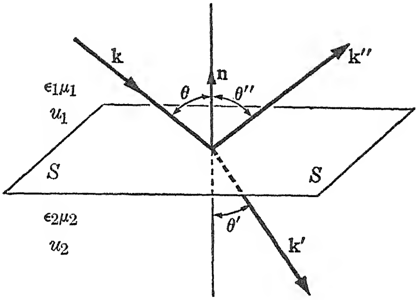


FIG. 11-2 The propagation vectors for refraction and reflection at a plane surface. **k** and **n**, the normal to the reflecting surface, determine the plane of incidence.

incident fields are represented by

$$\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)},\qquad \mathbf{H} = \frac{\mathbf{k}\times\mathbf{E}}{\omega\mu_1}. \tag{11-40}$$

The boundary conditions cannot be satisfied by a single progressive wave, since the tangential components of both **E** and **H** must be continuous across the boundary. If we use primed symbols for the wave in medium 2, and double primes for the reflected wave, we may write

$$\mathbf{E}' = \mathbf{E}'_0 e^{i(\mathbf{k}'\cdot\mathbf{r}-\omega t)},\quad \mathbf{H}' = \frac{\mathbf{k}'\times\mathbf{E}'}{\omega\mu_2},$$
$$\mathbf{E}'' = \mathbf{E}''_0 e^{i(\mathbf{k}''\cdot\mathbf{r}-\omega t)},\quad \mathbf{H}'' = \frac{\mathbf{k}''\times\mathbf{E}}{\omega\mu_1}. \tag{11-41}$$

The continuity of the tangential fields is possible only if the exponentials are the same at the boundary for all three fields. Thus we were justified in assuming throughout Eqs. (11-41) that the frequency is unchanged in the transmitted and reflected waves, and have the further conditions that

$$\mathbf{k}\cdot\mathbf{r} = \mathbf{k}'\cdot\mathbf{r} = \mathbf{k}''\cdot\mathbf{r} \tag{11-42}$$

over the boundary surface. It is evident from Eq. (11-42) that all the propagation vectors are coplanar. If for convenience we put the origin of the coordinate vector **r** in the boundary plane specified by the unit vector **n**, so that $\mathbf{n}\cdot\mathbf{r} = 0$ is the equation of the plane, then

$$\mathbf{n}\times(\mathbf{n}\times\mathbf{r}) = (\mathbf{n}\cdot\mathbf{r})\mathbf{n} - \mathbf{r} = -\mathbf{r}$$

on the boundary, and substitution in Eqs. (11-42) leads to

$$(\mathbf{k}-\mathbf{k}'')\times\mathbf{n}\cdot(\mathbf{n}\times\mathbf{r}) = 0,$$
$$(\mathbf{k}-\mathbf{k}')\times\mathbf{n}\cdot(\mathbf{n}\times\mathbf{r}) = 0. \tag{11-43}$$

Now the propagation vectors **k** and **k**$''$ in the same medium are equal in magnitude, and $|\mathbf{k}|/|\mathbf{k}'| = u_2/u_1$, the ratio of the phase velocities in the two media. Therefore the laws of reflection and refraction are implicit in Eq. (11–43), and in terms of the angles may be written in the usual forms:

$$\theta = \theta'',$$
$$\frac{\sin\theta}{\sin\theta'} = \frac{u_1}{u_2}. \tag{11–44}$$

The second of these equations is known as Snell's law.

The relations between the various amplitudes can be determined to satisfy the boundary conditions

$$\mathbf{n} \times (\mathbf{E} + \mathbf{E}'') = \mathbf{n} \times \mathbf{E}',$$
$$\mathbf{n} \times (\mathbf{k} \times \mathbf{E} + \mathbf{k}'' \times \mathbf{E}'')/\mu_1 = \mathbf{n} \times (\mathbf{k}' \times \mathbf{E}')/\mu_2,$$

where we have written the magnetic amplitudes in terms of those of the electric fields. It will suffice for us to consider the case in which $\mu_1 = \mu_2$. The orientation of **E** is arbitrary, but it may always be written as the sum of two components at right angles, and any difference in phase can be taken care of by making **E** complex.

(a) **E** *at right angles to the plane of incidence.* If **E** is perpendicular to the plane of incidence, all the electric field vectors are tangential to the surface, and

$$E_0 + E_0'' = E_0'. \tag{11–45}$$

Since $\mathbf{n} \cdot \mathbf{E}$ vanishes the expansion of the vector triple product in the second boundary condition leads to

$$E_0(\mathbf{n} \cdot \mathbf{k}) + E_0''(\mathbf{n} \cdot \mathbf{k}'') = E_0'(\mathbf{n} \cdot \mathbf{k}')$$

or, in terms of the angles of Fig. 11–2,

$$E_0 \cos\theta - E_0'' \cos\theta'' = \frac{u_1}{u_2} E_0' \cos\theta'. \tag{11–46}$$

Equations (11–45) and (11–46) may be combined to give the transmitted and reflected amplitudes in terms of that of the incident wave:

$$E' = E\frac{2\cos\theta}{\cos\theta + \dfrac{u_1}{u_2}\cos\theta'},$$
$$E'' = -E\frac{\dfrac{u_1}{u_2}\cos\theta' - \cos\theta}{\dfrac{u_1}{u_2}\cos\theta' + \cos\theta}. \tag{11–47}$$

Here use has been made of the fact that $\theta'' = \theta$, and the angle θ' may be eliminated by the use of Snell's law of refraction. $u_1/u_2 = n_{21}$, the relative index of refraction of medium 2 with respect to medium 1. n without subscripts is usually interpreted as n_{21}.

(b) **E** *in the plane of incidence.* In this case **H** is parallel to the boundary, and the derivation is formally identical with that above except that the relation

$$\mathbf{E} = -\mu\omega \frac{\mathbf{k} \times \mathbf{H}}{k^2}, \tag{11-48}$$

which follows from Maxwell's equations for a plane wave, leads to a slightly different distribution of constants. If we again assume that the two media have the same permeability, the final results are

$$\begin{aligned} H' &= H \frac{2 \cos \theta}{\cos \theta + \dfrac{u_2}{u_1} \cos \theta'}, \\ H'' &= -H \frac{\dfrac{u_2}{u_1} \cos \theta' - \cos \theta}{\dfrac{u_2}{u_1} \cos \theta' + \cos \theta}, \end{aligned} \tag{11-49}$$

for **E** in the plane of incidence.

Equations (11-47) and (11-49) are known as Fresnel's formulas, and were originally derived on the assumption of an elastic ether. Their transformation to other convenient forms, and the derivation of the resulting reflection and transmission coefficients, are left to the problems.

There are several interesting consequences of the fact that the two components are not transmitted and reflected equally. For the first case treated above, in which the electric vector is normal to the plane of incidence, there is reflection unless $\theta = \theta'$, i.e., unless the two media have the same optical properties. The reflected amplitude of Eq. (11-49), however, vanishes if $\theta + \theta' = 90°$, i.e., if $\mathbf{k}'$ and $\mathbf{k}''$ are perpendicular to each other.* The angle of incidence that satisfies this condition is called Brewster's angle, or the polarizing angle, and is defined in terms of Snell's law by

$$\tan \theta_0 = u_1/u_2. \tag{11-50}$$

Since at this angle the reflected beam consists entirely of the component whose electric vector is at right angles to the plane of incidence, a plate or a stack of plates of dielectric may be used to produce a plane polarized wave from unpolarized light.

---

* This is obvious from the formulas of Exercise 2, below.

On the other hand, there may be a whole range of angles for which there is no transmission. If u_2 is greater than u_1 the law of refraction (Snell's law) leads to values of $\sin\theta'$ that are greater than unity when θ exceeds $\sin^{-1}(u_1/u_2)$. Under these conditions the reflection is said to be total, but since the wave amplitudes may be complex in any case, we may apply Fresnel's formulas to examine the process in detail. If we eliminate θ' by means of Snell's law, using $n = u_1/u_2$, $(n < 1)$, the reflected amplitude for case (a), **E** normal to the plane of incidence, becomes

$$E'' = E\frac{\sin\theta\cos\theta - \sin\theta'\cos\theta'}{\sin\theta\cos\theta + \sin\theta'\cos\theta'}$$

$$= E\frac{\cos\theta - \dfrac{i}{n^2}\sqrt{\sin^2\theta - n^2}}{\cos\theta + \dfrac{i}{n^2}\sqrt{\sin^2\theta - n^2}}, \tag{11-51}$$

where the fact that $\cos\theta' = i\sqrt{\sin^2\theta - n^2}/n$ is imaginary has been made explicit. It is evident from Eq. (11-51) that the reflected amplitude is equal in magnitude to that of the incident wave, but the phase is changed:

$$E'' = Ee^{-2i\phi}, \tag{11-52}$$

where

$$\phi = \arctan\frac{\sqrt{\sin^2\theta - n^2}}{n^2\cos\theta}. \tag{11-53}$$

For the component whose electric vector is in the plane of incidence, we have from Eq. (11-49)

$$H'' = H\frac{\cos\theta - i\sqrt{\sin^2\theta - n^2}}{\cos\theta + i\sqrt{\sin^2\theta - n^2}}$$

$$= He^{-2i\psi}, \tag{11-54}$$

where

$$\psi = \arctan\frac{\sqrt{\sin^2\theta - n^2}}{\cos\theta}.$$

Thus the components are changed in phase by unequal amounts, and as a result plane polarized radiation is totally reflected with elliptical polarization. The phase difference between the two components, $2(\phi - \psi)$, depends on both the angle of incidence and the relative index of refraction. A simple computation leads to

$$\tan(\phi - \psi) = \frac{\cos\theta\sqrt{\sin^2\theta - n^2}}{\sin^2\theta}. \tag{11-55}$$

Although the transmission coefficient is zero for total reflection, the instantaneous fields do not vanish in the second medium. The fact that $\cos\theta'$ is a pure imaginary yields an exponential decrease of the field strength within the medium. The field in the second medium can be detected if an additional transparent object of large index is placed very near the surface.

11-6 Waves in conducting media and metallic reflection. The solutions of the general wave equation, Eq. (11-15), can be written in the same form as those of the propagation equation, Eq. (11-16), if we again assume a sinusoidal time dependence. We may write for any transverse component of **E**

$$\frac{\partial^2 \mathrm{E}}{\partial \zeta^2} = -(\mu\epsilon\omega^2 \mathrm{E} + i\omega\mu\sigma \mathrm{E}) = -K^2 \mathrm{E}, \tag{11-56}$$

where

$$K = \alpha + i\beta = k\sqrt{1 + \frac{i\sigma}{\epsilon\omega}}. \tag{11-57}$$

When Eq. (11-57) is solved for the real and imaginary parts of K, the results are

$$\alpha = k\sqrt{\frac{\sqrt{1 + \left(\frac{\sigma}{\epsilon\omega}\right)^2} + 1}{2}},$$

$$\beta = k\sqrt{\frac{\sqrt{1 + \left(\frac{\sigma}{\epsilon\omega}\right)^2} - 1}{2}}. \tag{11-58}$$

The fields are therefore

$$\mathbf{E} = \mathbf{E}_0 e^{-\beta\zeta} e^{i(\alpha\zeta - \omega t)},$$

$$\mathbf{H} = \frac{\mathbf{k} \times \mathbf{E}}{k\mu\omega}(\alpha + i\beta), \tag{11-59}$$

where **k** is along ζ, as before. Two features of Eq. (11-59) are immediately evident: the conductivity gives rise to exponential damping of the wave, and the electric and magnetic fields are no longer in phase.

The application of the Fresnel formulas to find the reflection at a plane metallic boundary is straightforward, but leads in general to rather complicated results. It is useful to consider the approximation, valid for all frequencies below the optical region, in which $\epsilon\omega \ll \sigma$. In that case,

$$K = (1 + i)k\sqrt{\frac{\sigma}{2\omega\epsilon}} = (1 + i)\sqrt{\frac{\mu\sigma\omega}{2}}. \qquad (11\text{–}60)$$

For simplicity, let a plane wave in a dielectric be incident normally on the surface of a conductor in the direction of ζ. The fields inside the conductor are then

$$\mathbf{E}' = \mathbf{E}_0' e^{-\sqrt{\frac{\mu\sigma\omega}{2}}\zeta} e^{i\left(\sqrt{\frac{\mu\sigma\omega}{2}}\zeta - \omega t\right)},$$
$$(11\text{–}61)$$
$$\mathbf{H}' = \frac{\mathbf{B}'}{\mu} = \frac{1 + i}{\sqrt{2}}\sqrt{\frac{\sigma}{\mu\omega}}\,\mathbf{n} \times \mathbf{E},$$

where we have measured ζ from the boundary surface. Both fields decrease with penetration, falling to $1/e$ of their surface values in a distance

$$\delta = \sqrt{\frac{2}{\mu\sigma\omega}}, \qquad (11\text{–}62)$$

which is known as the skin depth. Note that δ goes to zero as the conductivity approaches infinity, and is smaller for high frequencies. Equation (11–62) cannot be extrapolated to infinite frequencies, however, and an expression derived from Eq. (11–58) must be used where the approximation $\epsilon\omega \ll \sigma$ is not valid.

Inside the metal the energy density is no longer shared equally by the two fields: the time average of the electrical energy density is now $\epsilon\overline{E^2}/2$, while that associated with the magnetic fields is $\sigma\overline{E^2}/2\omega$, so that the ratio is $\epsilon\omega/\sigma$, which we have assumed is a very small number. Within the conductor the electrical energy density is negligible in comparison with the magnetic energy, and good conductors are essentially not penetrated by electric fields satisfying the condition that $\epsilon\omega \ll \sigma$.

The "reflecting power" of the metallic surface is readily found if we remember that the reflected wave travels toward increasing negative ζ, so that

$$\mathbf{E}'' = \mathbf{E}_0'' e^{-i(k\zeta + \omega t)},$$

and then impose the conditions that **E** and **H** are both continuous across the boundary. The ratio of the reflected energy to the incident energy is

$$\left|\frac{\mathbf{E}''}{\mathbf{E}}\right|^2 = 1 - 2\sqrt{\frac{2\epsilon_0\omega}{\sigma}} \qquad (11\text{–}63)$$

if higher powers of $\sqrt{2\epsilon_0\omega/\sigma}$ than the first are neglected.

11–7 General solutions of the homogeneous wave equation. In accord with the theorems of Fourier analysis, the general solution of the propagation equation in three dimensions can be expanded in terms of Fourier integrals over the three components of the wave propagation vector **k**, which is related to the frequency by the equation $k^2 = \omega^2/u^2$. The Fourier integrals are taken over all positive and negative values of each component of **k**, and in addition the resultant fields are summed over all possible polarizations. This gives the general expansion of the solution of the homogeneous wave equation as a superposition of plane waves. The form of the expansion is

$$\mathbf{E} = \sum_{\alpha=1}^{3} \left[\boldsymbol{\varepsilon}_\alpha \int_{-\infty}^{\infty} dk_1 \int_{-\infty}^{\infty} dk_2 \int_{-\infty}^{\infty} dk_3 a_{\mathbf{k}}^{\alpha} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \right], \qquad (11\text{–}64)$$

where the $\boldsymbol{\varepsilon}_\alpha$'s are unit vectors in the three coordinate directions. The $a_{\mathbf{k}}$ are amplitude functions of the frequency, and may be complex to give arbitrary starting phases. Such an expansion is often useful in the general theory of radiation. The "transversality condition" $\nabla \cdot \mathbf{E} = 0$ reduces the number of values of α from three to two.

We have seen in the previous section that good conductors effectively confine radiation fields of lower than optical frequencies to regions exterior to the conductors. For vacuum regions or dielectrics enclosed by conductors the wave equation becomes a homogeneous Sturm-Liouville equation with discrete eigenvalues, i.e., only certain values of the frequency parameter are consistent with the boundary conditions and Eq. (11–64) reduces to a sum over k. In Chapter 12 we shall investigate the general principles involved in such cases.

It is also possible to generalize a solution of the homogeneous wave equation in terms of an expansion in spherical waves, or in cylindrical waves, for either continuous or discrete ranges of k. Which of these various expansions is the more convenient to use depends on the symmetry properties of the problem under consideration, as we shall see.

SUGGESTED REFERENCES

Plane electromagnetic waves are treated in many books on electrodynamics and physical optics, and in general textbooks on theoretical physics. Our list includes some of the better and more easily available accounts.

M. Abraham and R. Becker, *Electricity and Magnetism*, Chapter X. Characteristically clear, rigorous, and rather elementary.

M. Born, *Optik*, Chapter 1. Probably the most complete and rigorous of the physical optics treatments.

J. A. Stratton, *Electromagnetic Theory*, Chapter V. A very useful account, presented in both detail and generality.

Briefer accounts will be found in:

J. H. Jeans, *The Mathematical Theory of Electricity and Magnetism.*

G. Joos, *Theoretical Physics.* Chapter XIX includes propagation in anisotropic media.

J. C. Slater and N. H. Frank, *Electromagnetism.*

W. R. Smythe, *Static and Dynamic Electricity.*

H. A. Lorentz, *The Theory of Electrons,* contains an excellent account of radiation pressure.

EXERCISES

1. Prove that the time average of the real part of the product of two complex vectors, both of which depend on the time as $e^{-i\omega t}$, is given by one-half the real part of the product of one with the complex conjugate of the other.

2. Prove that for the electric vector perpendicular to the plane of incidence Fresnel's formulas are

$$\mathbf{E}' = \frac{2\cos\theta\cos\theta'}{\sin(\theta+\theta')}\mathbf{E} \quad \text{for the transmitted wave,}$$

and

$$\mathbf{E}'' = \frac{\sin(\theta'-\theta)}{\sin(\theta'+\theta)}\mathbf{E} \quad \text{for the reflected wave.}$$

Prove also that when **E** is in the plane of incidence, the reflection is given by

$$\frac{|\mathbf{E}''|}{|\mathbf{E}|} = \frac{\tan(\theta-\theta')}{\tan(\theta+\theta')}.$$

3. By considering the waves with Fresnel amplitudes in the second medium, prove that the transmission coefficient vanishes under the conditions of total reflection.

4. Prove that for normal incidence the reflection coefficient, or percent of light reflected from the surface between two dielectrics, is given by $(n-1)^2/(n+1)^2$, where n is the relative index of refraction of medium 2 with respect to medium 1. What is the transmission coefficient? Compute the amount of light transmitted by a transparent glass plate in air at normal incidence, in terms of the index of refraction of the glass.

5. A plane wave is specularly reflected by a surface at an angle θ from the normal. Find the radiation pressure.

6. Estimate the ratio of the resistance of a conducting wire for high-frequency alternating currents to its direct current resistance. The result is known as Rayleigh's resistance formula.

7. Plane monochromatic waves are propagated parallel to the z-axis in both the positive and negative directions. At the origin the field strengths are given by

$$E_x = A\cos\omega t, \qquad E_y = 0,$$
$$H_x = 0, \qquad H_y = B\cos\omega t$$

Calculate the mean intensity of the radiation in each of the two directions in terms of A, B, and the constants of the medium.

CHAPTER 12

WAVE SOLUTIONS IN THE PRESENCE OF METALLIC BOUNDARIES

Before undertaking the theory of radiation, which examines the origin of electromagnetic waves from their sources, we shall consider further the solution of the homogeneous equation in the presence of conducting boundaries. The boundary value problem has been of great theoretical significance, and has many practical applications, particularly in the microwave region. We shall here be concerned primarily with the underlying principles, not the devices based on these principles. We may remark at the outset that there is no essential advantage in introducing potentials in the absence of charges and currents: it is equally convenient to work with the electric or magnetic field directly, remembering, however, that the two fields are not independent, but that one may be derived from the other by means of the field equations.

12–1 The nature of metallic boundary conditions. Let us review the general boundary conditions on the field vectors at a surface between medium 1 and medium 2:

$$\begin{aligned} \mathbf{n} \cdot (\mathbf{D}_1 - \mathbf{D}_2) &= \tau, \\ \mathbf{n} \times (\mathbf{E}_1 - \mathbf{E}_2) &= 0, \\ \mathbf{n} \cdot (\mathbf{B}_1 - \mathbf{B}_2) &= 0, \\ \mathbf{n} \times (\mathbf{H}_1 - \mathbf{H}_2) &= \mathbf{K}, \end{aligned} \tag{12–1}$$

where τ is used for the surface charge density to avoid confusion with the conductivity, and **K** is the surface current density. As usual, **n** is a unit vector normal to the surface. We have seen in Section 11–6 that for normal incidence an electromagnetic wave falls off very rapidly inside the surface of a good conductor, and that at the surface the tangential component of **H** may be large, while that of **E** vanishes in the limit of perfect conductivity. Let us examine the behavior of the normal components.

Consider a good conductor for which $\sigma/\omega\epsilon \gg 1$. The component of $\nabla \times \mathbf{H}$ normal to the surface is

$$\begin{aligned} \mathbf{n} \cdot (\nabla \times \mathbf{H}) &= -i\omega\epsilon\mathbf{n} \cdot \mathbf{E}, \text{ outside}, \\ \mathbf{n} \cdot (\nabla \times \mathbf{H}) &= (-i\omega\mathbf{D} + \sigma\mathbf{E}) \cdot \mathbf{n}, \text{ inside}. \end{aligned} \tag{12–2}$$

This component of $\nabla \times \mathbf{H}$ is conserved, since $\mathbf{n}\cdot(\nabla \times \mathbf{H}) = -\nabla\cdot(\mathbf{n} \times \mathbf{H})$. If we neglect the displacement current inside the metal,

$$\mathbf{n}\cdot\mathbf{E}\text{ (inside)} = -\frac{i\omega\epsilon}{\sigma}\mathbf{n}\cdot\mathbf{E}\text{ (outside)},$$

and there is, even in this approximation, a discontinuity in the normal component of **E**. Within the conductor the normal component of the electric field, as well as its tangential component, becomes vanishingly small as the conductivity approaches infinity. The normal component of **B**, on the other hand, suffers no discontinuity at the surface.

If **E** vanishes inside a perfect conductor the curl of **E** also vanishes, and the time rate of change of **B** is correspondingly zero. This means that there are no oscillatory fields whatever inside such a conductor, and that the boundary values of the fields outside are given by

$$\begin{aligned} \mathbf{n}\cdot\mathbf{D} &= \tau, \\ \mathbf{n}\times\mathbf{E} &= 0, \\ \mathbf{n}\cdot\mathbf{B} &= 0, \\ \mathbf{n}\times\mathbf{H} &= \mathbf{K}. \end{aligned} \tag{12–3}$$

Thus the electric field is normal and the magnetic field tangential at the surface of a perfect conductor. For good conductors these conditions lead to excellent representations of the geometrical configurations of external fields, but they correspond to complete neglect of some important features of real fields, such as losses in cavities and attenuation of signal in the case of transmitted waves.

It is useful for the estimation of losses to see how the tangential and normal field components compare with the first approximation in $1/\sigma$. The tangential component of **E** is strictly continuous, so that we have, from Section 11–6,

$$\mathbf{H} = \frac{1+i}{\sqrt{2}}\sqrt{\frac{\sigma}{\mu\omega}}\,\mathbf{n}\times\mathbf{E} \tag{12–4}$$

at the surface of the conductor. Let us assume, without obtaining a complete solution, that a wave with **H** tangential and **E** normal is propagated parallel to the surface of the metal. If k is the component of the propagation vector along the surface, then just outside the surface,

$$|\mathbf{H}_{\parallel}| = \frac{k}{\mu\omega}|\mathbf{E}_{\perp}|. \tag{12–5}$$

But a tangential component of **H** is accompanied by a small tangential component of **E**, according to Eq. (12–4). By comparing these two expressions, we see that

$$\frac{|\mathbf{E}_{\|}|}{|\mathbf{E}_{\perp}|} \simeq k\sqrt{\frac{2}{\mu\omega\sigma}} = \frac{\delta}{\bar{\lambda}}, \tag{12–6}$$

and the ratio of the tangential component of **E** to its normal component has the same order of magnitude as the skin depth divided by the wavelength. It can be readily shown that the ratio of the normal component of **H** to its tangential component is of this same magnitude, but for the estimation of losses it is the tangential component of **E** in which we shall be interested. Thus we see that in the limit of high conductivity, which means vanishing skin depth, no fields penetrate the conductor, and the boundary conditions are just those given by Eqs. (12–2). We shall investigate the solution of the homogeneous wave equation subject to such boundaries.

12–2 Eigenfunctions and eigenvalues of the wave equation. The vector wave equation in Cartesian coordinates is simply a set of three scalar equations, one in each of the rectangular components of the vector. In other coordinate systems the situation is more complicated, but we shall see that at least in cylindrical and spherical polar coordinates the fields may also be written in terms of scalar functions. Our problem is to solve an equation of the form

$$\nabla^2\Psi - \frac{1}{c^2}\frac{\partial^2\Psi}{\partial t^2} = 0, \tag{12–7}$$

where Ψ may represent a component of **E** or **H**, in such a way that certain boundary conditions are satisfied.

Equation (12–7), unlike Laplace's equation, is not symmetrical in its independent variables: the second derivative with respect to the time occurs with the sign opposite to that of the space derivatives. The resulting character of possible solutions corresponds to a fundamental physical distinction from the case of static potentials: entirely oscillatory solutions are now allowed in a shielded region free of sources, whereas, for example, the static field inside a grounded conductor is zero throughout.

We may assume simple harmonic time dependence for the function $T(t)$ in the expression

$$\Psi(x,y,z,t) = \psi(x,y,z)T(t) = \psi(x,y,z)e^{-i\omega t}, \tag{12–8}$$

where ψ is a function of the space variables alone, and satisfies the differential equation

$$\nabla^2\psi = -\frac{\omega^2}{c^2}\psi = -k^2\psi. \tag{12–9}$$

GRICULTURAL

In rectangular coordinates, we may set $\psi = X(x)Y(y)Z(z)$, whereupon separation yields

$$\begin{aligned} X'' + k_1^2 X &= 0, \\ Y'' + k_2^2 Y &= 0, \\ Z'' + k_3^2 Z &= 0, \end{aligned} \tag{12-10}$$

with

$$k_1^2 + k_2^2 + k_3^2 = k^2 = \omega^2/c^2. \tag{12-11}$$

Equation (12–11) indicates that there are no more than the expected number of separation constants, namely three, corresponding to the four independent variables. Equations (12–10), or their counterparts in other coordinate systems, must be solved subject to the boundary conditions on the field component represented by or related to the scalar function ψ.

Let us examine this problem somewhat more generally. Each of Eqs. (12–10) is a simple example of the general form

$$\frac{d}{dx}\left[p(x)\frac{dX}{dx}\right] + [q(x) + \lambda r(x)]X = 0, \tag{12-12}$$

where λ is a parameter, in our case the separation constant. The solutions are required to be continuous and bounded throughout an interval $a \leq x \leq b$, and to satisfy certain boundary conditions at $x = a$, $x = b$. [As to the nature of the unspecified functions $p(x)$, $q(x)$, $r(x)$, we need only assume that they are such as to introduce no singularities within the interval, although there may be singularities at the boundaries.] The boundary conditions can be satisfied in general only by certain values of λ, and by particular solutions of Eq. (12–12) corresponding to these values of λ. The allowed values of λ are called eigenvalues, and the allowed solutions are called eigenfunctions. It will be recognized that we have met this one-dimensional problem before, in applying the method of separation of variables to the solution of static potential problems, but here the solution of the entire three-dimensional problem as a whole may be characterized by an eigenvalue related to the frequency.

To facilitate discussion at first, let us consider the case where $q(x) = 0$, $r(x) = 1$, under the conditions that $X(x) = 0$ at the ends of the interval. Equation (12–12) becomes

$$\frac{d}{dx}\left[p(x)\frac{dX}{dx}\right] = -\lambda X. \tag{12-13}$$

From Eq. (12–13) it can be seen that if we have a solution $X(x)$ equal to zero at $x = a$, it can never satisfy the same condition at $x = b$ if λ is negative, for in that case it would continue to depart from its zero value once it began to do so. For positive values of λ, however, the curvature is opposite in sign to that of $X(x)$, and for some value of λ the function will be brought back to the x-axis just at b. Further increase in λ will cause $X(x)$ to change sign short of b, but a value will be reached for which $X(x)$ will again be zero at $x = b$, this time with one zero in the interval $a \leq x \leq b$. Thus we see that there is an absolute lower bound for λ, and the eigenvalues may be numbered $\lambda_0, \lambda_1, \cdots$, with subscripts corresponding to the number of zeros in the interval.

This whole argument may be repeated for the more general case of Eq. (12–12), except that the sign of the curvature will depend on $q(x)$ and $r(x)$. There is still a lowest allowed value of λ, but the curvature is determined by the sign of the entire coefficient of $X(x)$ in the differential equation. For example, Bessel's equation becomes

$$\frac{d}{dr}\left(r\frac{dJ_n}{dr}\right) + \left(k^2r - \frac{n^2}{r}\right)J_n = 0$$

when written in this form, and the curvature will not be such as to bring the function back toward the axis unless $k^2r > n^2/r$.

If the boundary conditions are such that either $X(x)$ or its first derivative must vanish at the ends of the interval, the eigenfunctions of Eq. (12–12) are orthogonal over the interval. Let X_n and X_m be two allowed solutions corresponding to two eigenvalues of the parameter, λ_n and λ_m. To prove orthogonality, let the equation for X_n be multiplied by X_m, that for X_m by X_n, one subtracted from the other, and the whole integrated over the interval. Since either all functions X or their derivatives vanish at the boundaries, the result is

$$(\lambda_m - \lambda_n)\int_a^b r(x)X_mX_n\,dx = 0. \tag{12–14}$$

The function $r(x)$ is often called the "weight function" in mathematical treatments.

If more than one solution is allowed for a particular value of λ the problem is said to be degenerate, and the number of independent solutions corresponding to that value of λ is called the degree of the degeneracy. In this case Eq. (12–14) fails to show orthogonality, i.e., λ_m may be equal to λ_n for $n \neq m$. The solutions may nevertheless be chosen to be orthogonal, due to the fact that any linear combination of the independent solutions belonging to a given λ is also a solution, and these linear combinations may be made in such a way that they are orthogonal to each other. The choice of coefficients in the orthogonal linear combinations is not unique, but is usually made so that the final solutions correspond to some readily identifiable physical property of the system. In any case, we may always speak of the independent solutions as orthogonal, assuming that orthogonalization has been carried out for degenerate λ's.

In any coordinate system in which the wave equation is separable the function ψ is a product of single variable functions such as $X(x)$. The frequency for which there is an allowed solution of Eq. (12–9) depends on three allowed parameters λ, one for each coordinate variable. It is instructive to prove that the wave equation generates orthogonal functions. If we apply Green's theorem to two product eigenfunctions, ψ_n and ψ_m, we obtain

$$\int (\psi_m \nabla^2 \psi_n - \psi_n \nabla^2 \psi_m)\, dv = \int (\psi_m \nabla \psi_n - \psi_n \nabla \psi_m) \cdot d\mathbf{S}.$$

Either boundary condition, the vanishing of the field component or its normal derivative, causes the surface integral to vanish. The volume integral can be written, in view of Eq. (12–9), as

$$(k_n^2 - k_m^2) \int \psi_m \psi_n \, dv = 0, \tag{12–15}$$

which is the desired orthogonality relation. By comparison of Eqs. (12–14) and (12–15) it is seen that the "weight function" $r(x)$ is just the factor required to give the correct volume element in three dimensions.

In closed cavities, then, with boundaries on which the field components or their derivatives vanish, we may expect to find characteristic electromagnetic vibrations, similar to the modes of a vibrating string. The frequencies of these vibrations will correspond to a lowest eigenvalue, the "fundamental," and an infinite discrete set of higher frequencies. These modes of vibration are sharp and independent insofar as the approximation of perfect conductivity is valid; in practice they will not sustain themselves indefinitely, and we shall have to look at the effect of finite conductivity in order to estimate losses. It should be pointed out that the actual configuration of fields in a cavity will depend on the method of excitation.

Although every one of the four single variable factors in $\Psi(x,y,z,t)$ may now be an oscillatory function, only three correspond to closed boundaries. The fourth independent variable, time, is still "open," and the complete solution will depend on the initial conditions, $\Psi(x,y,z,t_0)$ and $(\partial\Psi/\partial t)_{t=t_0}$, at some particular time $t = t_0$. There is no solution corresponding to the specification of both initial and final conditions.

12–3 Cavities with rectangular boundaries. Let us consider a vacuum region totally enclosed by rectangular conducting walls. In this case, all field components are products of the solutions of Eqs. (12–10), with k_1, k_2, and k_3 so chosen that the electric fields are normal to the walls at the boundary and the magnetic fields tangential. If A, B, and C are the dimensions of the cavity, it may be readily verified that the electric field components are

$$\begin{aligned} E_x &= E_1 \cos k_1x \sin k_2y \sin k_3z \, e^{-i\omega t}, \\ E_y &= E_2 \sin k_1x \cos k_2y \sin k_3z \, e^{-i\omega t}, \\ E_z &= E_3 \sin k_1x \sin k_2y \cos k_3z \, e^{-i\omega t}, \end{aligned} \tag{12–16}$$

where the eigenvalues of the separation parameters needed to satisfy the boundary conditions are

$$k_1 = l\pi/A, \quad k_2 = m\pi/B, \quad k_3 = n\pi/C, \tag{12–17}$$

with l, m, and n integers. The allowed frequencies are thus given by

$$\frac{\omega^2}{c^2} = k^2 = \pi^2\left(\frac{l^2}{A^2} + \frac{m^2}{B^2} + \frac{n^2}{C^2}\right). \tag{12–18}$$

It is clear from Eqs. (12–16) that at least two of the integers l, m, n must be different from zero in order to have nonvanishing fields. The magnetic fields obtained by the use of $\nabla \times \mathbf{E} = i\omega\mathbf{B}$ automatically satisfy the appropriate boundary conditions, and are seen to be in time quadrature with the electric fields. Thus the sum of the total electric and magnetic energies within the cavity is constant although the two terms fluctuate separately.

The amplitudes of the electric field components are not independent, but are related by the divergence condition $\nabla \cdot \mathbf{E} = 0$, which yields

$$E_1k_1 + E_2k_2 + E_3k_3 = 0. \tag{12–19}$$

If $\mathbf{k}$ is considered as a vector with components k_1, k_2, k_3, Eq. (12–19) is the condition that $\mathbf{E}$ is a vector perpendicular to $\mathbf{k}$. There are, in general, two linearly independent vectors $\mathbf{E}$ that satisfy this condition, corresponding to two possible polarizations. (The exception is the case that one of the integers l, m, n is zero, so that $\mathbf{E}$ is fixed in direction.) These vectors may be chosen arbitrarily so long as they are not coplanar, and each is accompanied by a magnetic field at right angles. Thus again $\mathbf{E}$, $\mathbf{H}$, and $\mathbf{k}$ form an orthogonal set of vectors. The fields corresponding to a given set of integers l, m, n constitute a mode of vibration such as we predicted in the previous section. The total field is a sum over all possible modes of vibration, with amplitude factors that depend on the method of excitation.

Actual conducting walls gradually absorb energy from the cavity at a rate that can be readily estimated from the considerations of Section 12–1. For finite σ the small tangential component of $\mathbf{E}$ may be found from Eq. (12–4)

$$\mathbf{E}_{\|} = \frac{1-i}{\sqrt{2}}\sqrt{\frac{\mu\omega}{\sigma}}\,\mathbf{H}_{\|} \times \mathbf{n}.$$

The tangential component of **H** will be slightly different from the ideal solution, but its relative change will be very small and may be neglected. To find the power loss, we may compute the time average of the Poynting vector into the wall at the surface,

$$\overline{\mathbf{N}} = \frac{\mathrm{Re}(\mathbf{E}_{\|} \times \mathbf{H}_{\|})}{2} = \frac{\mathbf{n}}{2}\sqrt{\frac{\mu\omega}{2\sigma}}\,\mathbf{H}_{\|0}^{2} = \frac{\mathbf{H}_{\|0}^{2}}{2\sigma\delta}. \tag{12–20}$$

Here $\mathbf{H}_{\|0}$ is the crest value of the tangential magnetic field, which equals the crest value of the surface current **K**. If we define a surface resistance such that

$$\overline{\mathbf{N}} = \mathbf{n}(\overline{\mathbf{K}^2}R_s),$$

then

$$R_s = \frac{1}{\sigma\delta}.$$

The total power loss in a cavity is obtained if we integrate Eq. (12–20) over the entire area of the walls. Mode losses are often expressed by giving the "Q" for a cavity, defined by

$$Q = \frac{\text{energy stored in the cavity}}{\text{energy lost per cycle to the walls}},$$

which is independent of the amplitude of mode excitation.

12–4 Cylindrical cavities. The rectangular parallelepiped we have considered is a special kind of right cylinder, and has many features in common with cylindrical cavities of arbitrary cross section. In every cavity the allowed values of k, and thus the allowed frequencies, are determined by the geometry of the cavity. But we have seen that for each set of k_1, k_2, k_3 in the rectangular cavity there are, in general, two linearly independent modes, i.e., the polarization remains arbitrary. We may take advantage of this arbitrariness to classify modes into two kinds according to the orientation of the field vectors: let us choose one mode such that the electric field lies in the cross-sectional plane, and the other so that the magnetic vector lies in this plane. This classification into transverse electric (TE) and transverse magnetic (TM) modes turns out to be possible for all cylindrical cavities, although the rectangular parallelepiped with unequal axes is unique in having one mode of each kind correspond to the same allowed frequency. (There is, of course, a higher degree of degeneracy if any axes of the parallelepiped are equal.)

To show that this classification is both possible and complete for the modes of any cylindrical cavity, we first note that the factor depending on the altitude coordinate z may be separated out of the wave equation. The boundary conditions at $z = 0$ and $z = C$ demand that the z dependence be given by either $\sin k_3 z$ or $\cos k_3 z$, where $k_3 = n\pi/C$. In other words, every field component satisfies the equation

$$\left(\frac{\partial^2}{\partial z^2} + k_3^2\right)\begin{Bmatrix}\mathbf{E}\\ \mathbf{H}\end{Bmatrix} = 0 \tag{12-21}$$

as well as

$$(\nabla^2 + k^2)\begin{Bmatrix}\mathbf{E}\\ \mathbf{H}\end{Bmatrix} = 0, \tag{12-22}$$

with k as yet undetermined. The field equations

$$\begin{aligned}\nabla \times \mathbf{E} &= i\omega\mu\mathbf{H},\\ \nabla \times \mathbf{H} &= -i\omega\epsilon\mathbf{E}\end{aligned} \tag{12-23}$$

must also be satisfied. If we write each vector and each operator in Eqs. (12-23) as the sum of a transverse part, designated by the subscript s, and a component along z, we find for the transverse fields

$$\begin{aligned}i\omega\mu\mathbf{H}_s &= \nabla_s \times \mathbf{E}_z + \nabla_z \times \mathbf{E}_s,\\ -i\omega\epsilon\mathbf{E}_s &= \nabla_s \times \mathbf{H}_z + \nabla_z \times \mathbf{H}_s,\end{aligned} \tag{12-24}$$

since, for example, $\nabla_s \times \mathbf{E}_s = i\omega\mu\mathbf{H}_z$. When one of Eqs. (12-24) is substituted for the transverse field on the right side of the other, and use is made of Eq. (12-21), there results

$$\begin{aligned}\mathbf{E}_s &= \frac{\nabla_s \dfrac{\partial E_z}{\partial z}}{k^2 - k_3^2} + \frac{i\omega\mu}{k^2 - k_3^2}\nabla_s \times \mathbf{H}_z,\\ \mathbf{H}_s &= \frac{\nabla_s \dfrac{\partial H_z}{\partial z}}{k^2 - k_3^2} - \frac{i\omega\epsilon}{k^2 - k_3^2}\nabla_s \times \mathbf{E}_z.\end{aligned} \tag{12-25}$$

All transverse fields are thus expressed in terms of the z-components, each of which satisfies the differential equation

$$[\nabla_s^2 + (k^2 - k_3^2)]\begin{Bmatrix}E_z\\ H_z\end{Bmatrix} = 0, \tag{12-26}$$

where ∇_s^2 is the two-dimensional Laplacian operator.

The conditions on E_z and H_z at the boundary of the cylinder cross section differ from each other, however: E_z must vanish on the boundary, while the normal derivative of H_z must vanish to satisfy the conditions on the second of Eqs. (12-25). When the cross section is a rectangle, these two conditions may lead to the same eigenvalues of $(k^2 - k_3^2) = k_s^2 = k_1^2 + k_2^2$, as we have seen. Otherwise they correspond to two different frequencies, one for which E_z is permitted but $H_z = 0$, the other with H_z but no E_z. In every case, it is possible to classify the modes as transverse magnetic or transverse electric.

The frequencies are determined by the eigenvalues of Eqs. (12–21) and (12–26). If we denote the functional dependence of E_z or H_z on the plane cross section coordinates by $f(x,y)$, we may write Eq. (12–26) as

$$\nabla_s^2 f = -(k^2 - k_3^2)f = -k_s^2 f. \tag{12–27}$$

Let us first show that $k_s^2 > 0$, and hence $k > k_3$. From Green's theorem, we note that

$$-k_s^2 \int f^2\, dv + \int (\nabla_s f)^2\, dv = \int f\, \nabla f \cdot d\mathbf{S}.$$

If either f or its normal derivative is to vanish on S, the cylindrical surface, then

$$k_s^2 = \frac{\int (\nabla_s f)^2\, dv}{\int f^2\, dv} > 0. \tag{12–28}$$

We have already seen that $k_3 = n\pi/C$. The allowed values of k_s depend both on the geometry of the cross section and the nature of the mode.

For TM modes $H_z = 0$, and the z dependence of E_z is given by $\cos n\pi z/C$. Equation (12–27) must be solved subject to the condition that f vanishes on the boundaries of the plane cross section, thus completing the determination of E_z and k. The transverse fields are special cases of Eqs. (12–25):

$$\begin{aligned} \mathbf{E}_s &= \frac{1}{k_s^2} \nabla_s \frac{\partial E_z}{\partial z}, \\ \mathbf{H} &= -\frac{i\omega\epsilon}{k_s^2} \nabla_s \times \mathbf{E}_z \\ &= \frac{i\omega\epsilon}{k_s^2} \hat{\mathbf{z}} \times \nabla_s E_z, \end{aligned} \tag{12–29}$$

where $\hat{\mathbf{z}}$ is a unit vector along the axis of the cylinder.

For TE modes in which $E_z = 0$, the condition that H_z vanish at the ends of the cylinder demands the use of $\sin n\pi z/C$, and k_s must be such that the normal derivative of H_z is zero at the walls. Equations (12–25), giving the transverse fields, then become

$$\mathbf{H}_s = \frac{1}{k_s^2}\nabla_s \frac{\partial H_z}{\partial z},$$

$$\mathbf{E} = \frac{i\omega\mu}{k_s^2}\nabla_s \times \mathbf{H}_z = -\frac{i\omega\mu}{k_s^2}\hat{\mathbf{z}} \times \nabla_s H_z, \tag{12-30}$$

and the mode determination is completed.

12-5 Circular cylindrical cavities. Let us apply the methods of the previous section to the TM modes of a right circular cylinder. We may write

$$E_z = Af(r,\varphi)\cos(k_3 z)e^{-i\omega t},$$

where $f(r,\varphi)$ satisfies the equation

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 f}{\partial \varphi^2} + k_s^2 f = 0. \tag{12-31}$$

Separation of variables leads to Bessel's equation for the radial coordinate, just as in electrostatic problems of cylindrical symmetry. If the axis $r = 0$ is included in the cavity, that solution must be chosen which is regular at the origin. For the TM mode in a cavity of radius r_0 and height C,

$$E_z = AJ_m(k_l r)e^{im\varphi}\cos(k_3 z)e^{-i\omega t}. \tag{12-32}$$

The k_l are determined by the condition that $J_m(k_l r_0) = 0$, and are the eigenvalues of k_s in Eq. (12-31). The frequency of any mode is given by

$$\frac{\omega^2}{c^2} = k^2 = k_l^2 + \frac{n^2\pi^2}{C^2}.$$

Knowledge of the zeros of the Bessel functions is thus necessary in order for us to ascertain the numerical values of the frequencies. If l is the ordinal number of a zero of a particular Bessel function of order m (l increases with increasing values of the argument) then each mode is characterized by three integers, l, m, n, as in the rectangular case. All transverse components of the fields may be found by taking the appropriate derivatives of E_z, as indicated in Eqs. (12-29).

The frequencies of the transverse electric modes are determined by the fact that $J_m(k_l r_0)$ must have zero slope at the boundary. The eigenvalues k_l are thus different from those of the TM case, and the cavity frequencies are correspondingly different. Again there is an infinite discrete set of modes, each characterized by three integers.

Sectors of cylinders can be treated by means of Bessel functions of fractional order, since m is no longer an integer but is determined by the boundary condition on the φ = constant surfaces. Cavities in the form of coaxial cylinders are amenable to the same methods of solution, except that the two boundary surfaces require the use of the general solution of Bessel's equation. Perhaps the most interesting feature of the double-walled cavity, however, is that in addition to TM and TE modes there are also TEM modes for which $E_z = 0 = H_z$. So far as these modes are concerned, the boundaries of the multiply connected region behave like internal sources for the fields. It is easy to see physically that there is even a zero-frequency mode having no electric field components, which would correspond to an upward current in the outside cylinder and a downward current in the inside one, for example.

12–6 Wave guides.

Let us consider the possibility of transmitting electromagnetic waves along the axis of a wave guide, which is simply a long cylinder with open ends. To represent a progressive wave along z, we may write the dependence of the fields on the coordinate variables and the time as

$$f(x,y)e^{i(k_g z-\omega t)}. \tag{12–33}$$

The "guide propagation constant," k_g, is just the k_3 of Eq. (12–21) except that it is no longer restricted by boundary conditions to discrete values. The general considerations of Section 12–4 apply, so that we may treat TM and TE modes separately.

One other case, briefly noted in Section 12–5, is of greater importance in this connection. If the boundaries are not singly connected, i.e., if the region is bounded by at least two conducting cylinders, the two-dimensional Laplacian of the fields, or $\nabla_s^2 f(x,y)$, can vanish subject to the boundary conditions. The solution of the two-dimensional electrostatic problem will then define the solution in a transverse plane and $k^2 = k_g^2$, so that the velocity of propagation in vacuum is simply c. Such a wave will have no field components along the axis of the cylinder: E_z vanishes on all boundaries and, being a solution of $\nabla_s^2 E_z = 0$, will vanish everywhere. Similarly, since the normal derivative of H_z vanishes on all boundaries and since $\nabla_s^2 H_z = 0$, H_z vanishes everywhere. Hence such a wave is purely transverse (TEM mode). The ordinary coaxial and bifilar line structures are typical examples of such cylinders.

If the boundary is singly connected, the solutions for f are identical with those for cylindrical cavities already discussed. Although k_g is not restricted in magnitude, we note that for every eigenvalue of the two-dimensional equation, k_s, there is a lowest value of k, namely, $k = k_s$ (often designated by k_c for wave guides), for which k_g is real. This corresponds to a "cutoff frequency" below which waves are not transmitted by that mode, and the fields fall off exponentially with increasing z. An absolute cutoff frequency is associated with the mode of lowest frequency, i.e., with the lowest value of k_s. In general,

$$k_s^2 = k^2 - k_g^2. \tag{12–34}$$

For real k_g it is obvious from the form of Eq. (12–33) that the wave is propagated along the guide with a phase velocity

$$v_p = \frac{\omega}{k_g} = \frac{k}{k_g}c.$$

Since a minimum condition for propagation is that k is greater than k_g, it is evident that the wave or phase velocity is greater than that of electromagnetic waves in free space. This velocity is not constant, however, but depends on the frequency. If we add two waves of slightly different frequencies, we obtain a factor

$$e^{i(k_gz-\omega t)}[1 + e^{i(\Delta k_gz-\Delta\omega t)}],$$

indicating that the original wave is modulated by the expression in brackets. The modulation itself is propagated with the speed

$$v_g = \frac{\Delta\omega}{\Delta k_g} \to \frac{d\omega}{dk_g}, \tag{12–35}$$

which is called the group velocity. For wave guides, it follows from Eq. (12–34) that

$$v_g = \frac{d\omega}{dk_g} = \frac{k_g}{k}c, \tag{12–36}$$

and v_g is thus always smaller than c. As a matter of fact,

$$v_pv_g = c^2.$$

Only if a pulse consists of a narrow range of frequencies can it be said to be transmitted with a definite velocity. The group velocity is usually evaluated at some median frequency.

The fields may be expressed even more simply in terms of the longitudinal field components than in the case of cylindrical cavities, since here $\partial/\partial z = ik_g$. It may be seen from Eqs. (12–25) that

for TM modes ($H_z = 0$),

$$\begin{aligned} \mathbf{E}_s &= \frac{ik_g}{k_s^2}\nabla_sE_z, \\ \mathbf{H}_s &= \frac{\omega\epsilon}{k_g}\hat{\mathbf{z}} \times \mathbf{E}_s \quad \text{or} \quad c\mathbf{B}_s = \frac{k}{k_g}\hat{\mathbf{z}} \times \mathbf{E}_s, \end{aligned} \tag{12–37}$$

and for TE modes ($E_z = 0$),

$$\mathbf{H}_s = \frac{ik_g}{k_s^2} \nabla_s H_z,$$

$$\mathbf{E}_s = -\frac{\omega\mu}{k_g} \hat{\mathbf{z}} \times \mathbf{H}_s \quad \text{or} \quad \mathbf{E}_s = \frac{ck}{k_g} \mathbf{B}_s \times \hat{\mathbf{z}}. \tag{12-38}$$

The time average z-component of the Poynting vector $\mathbf{N}$ is given by

$$\overline{\mathbf{N}_z} = \frac{|\mathbf{E}_s \times \mathbf{H}_s^*|}{2}$$

$$= \sqrt{\frac{\mu_0}{\epsilon_0}} \frac{k}{k_g} \frac{H_{s0}^2}{2} \quad \text{for TE modes}$$

$$= \sqrt{\frac{\mu_0}{\epsilon_0}} \frac{k_g}{k} \frac{H_{s0}^2}{2} \quad \text{for TM modes,} \tag{12-39}$$

where the subscript 0 denotes the crest value of the field. Wave guide losses may be estimated by integrating Eq. (12–20) over the walls of the guide for any given mode. The energy flow of a progressive wave will be attenuated as e^{-Kz}, where

$$K = \frac{\text{power loss per unit length of guide}}{\text{power transmitted}}$$

$$= \frac{1}{2\sigma\delta} \int (H_s^2 + H_z^2)\, dS \Big/ \tfrac{1}{2} \int Z_g H_s^2\, dS, \tag{12-40}$$

with the numerator integrated over unit length of wall and the denominator integrated over the guide cross section. The "guide impedance" Z_g is given by

$$Z_g = \sqrt{\frac{\mu_0}{\epsilon_0}} \frac{k}{k_g} \quad \text{for TE modes,}$$

$$Z_g = \sqrt{\frac{\mu_0}{\epsilon_0}} \frac{k_g}{k} \quad \text{for TM modes.}$$

The exponential attenuation that results when $k < k_s$ does not involve power loss in the walls, but is generated by a superposition of interfering reflections.

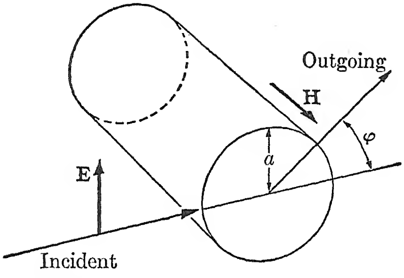


FIG. 12–1 Cross section of a conducting circular cylinder with radiation incident at right angles to the axis. **H** is parallel to the axis of the cylinder.

12–7 Scattering by a circular cylinder. The scattering of a plane wave by a cylindrical conductor offers an instructive application of boundary conditions on the homogeneous wave equation solutions. Consider a perfectly conducting cylinder of radius a, its axis coincident with the z-coordinate axis, as shown in Fig. 12–1. Let there be incident normal to the axis of the cylinder a plane wave of angular frequency ω, polarized in such a way that its magnetic vector is in the z-direction. The total field will consist of the incident wave and that scattered (reflected) by the cylinder. This field is essentially a transverse electric mode, with $k_3 = 0$ and no restrictions on the frequency.

The generating function for Bessel functions permits us to write down immediately the appropriate expansion for the magnetic vector of the incoming wave:

$$H_{z0}\, e^{i(kr\cos\varphi - \omega t)} = H_{z0} \sum_{m=-\infty}^{\infty} J_m(kr) i^m e^{i(m\varphi - \omega t)}. \tag{12–41}$$

In order to represent an outgoing wave from the scattering cylinder, we must choose that combination of Bessel functions which varies as $e^{ikr}/\sqrt{r}$ for large values of r. This combination is just $J_m(kr) + iN_m(kr)$, for which the standard notation is $H_m^{(1)}(kr)$; we may drop the superscript, and note that this function will not be confused with a magnetic field because of the subscript m and the explicit argument kr. Thus the scattered wave is of the form

$$H_{z0} \sum_m a_m H_m(kr) e^{i(m\varphi - \omega t)}. \tag{12–42}$$

The coefficients a_m are determined by the condition that the tangential component of the total electric field vanishes on the surface of the cylinder. Since H_z does not depend on z, it remains the only wave component of the magnetic field, a result we should also expect for physical reasons. It therefore follows from the curl relation that

$$E_\varphi = \frac{-i}{\epsilon_0 \omega} \frac{\partial}{\partial r} H_z. \tag{12–43}$$

AGRICULTURA

For each term of the series, therefore, the boundary condition requires that

$$i^m J'_m(ka) + a_m H'_m(ka) = 0,$$

where the prime denotes differentiation with respect to the argument, or with respect to r. Solving for a_m, we obtain

$$a_m = -i^m \frac{J'_m(ka)}{H'_m(ka)}. \tag{12–44}$$

We are interested in the scattered wave at large distances from the cylinder. For large values of its argument,

$$H_m(kr) \rightarrow \sqrt{\frac{2}{\pi kr}}\, e^{ikr} e^{-(2m+1)\frac{\pi}{4}i}, \tag{12–45}$$

and the scattered wave fields are thus

$$H_z = H_{z0} \sqrt{\frac{2}{\pi kr}}\, e^{i(kr-\omega t)} \Sigma a_m e^{i\left(m\varphi - \frac{2m+1}{4}\pi\right)}$$

$$E_\varphi = \sqrt{\frac{\mu_0}{\epsilon_0}}\, H_z,$$

if we neglect terms that fall off more rapidly with increasing r.

The Poynting vector associated with these fields is, of course, pointed along positive r. Its time average is

$$\overline{N}_r = \frac{1}{2} \operatorname{Re}\,(E_\varphi H_z^*) = \frac{1}{2} \sqrt{\frac{\mu_0}{\epsilon_0}}\, |H_z|^2. \tag{12–46}$$

If we integrate **N** over a cylindrical surface to determine the total rate of energy scattered per unit axial length of the cylinder, the double sum in the product reduces to a single sum because of the orthogonality of the functions $e^{im\varphi}$, and we obtain

$$\frac{dW}{dt} = \int_0^{2\pi} Nr\, d\varphi = \frac{1}{2} \sqrt{\frac{\mu_0}{\epsilon_0}}\, |H_{z0}|^2 \frac{4}{k} \sum_m \left| \frac{J'_m(ka)}{H'_m(ka)} \right|^2 \tag{12–47}$$

or, in terms of real Bessel functions,

$$\frac{dW}{dt} = \frac{1}{2} \sqrt{\frac{\mu_0}{\epsilon_0}}\, H_{z0}^2 \frac{4}{k} \sum \frac{[J'_m(ka)]^2}{[J'_m(ka)]^2 + [N'_m(ka)]^2}. \tag{12–47$'$}$$

The time average of the energy flow per unit area in the incident beam is

$$\overline{N_0} = \frac{1}{2}\sqrt{\frac{\mu_0}{\epsilon_0}}\,H_{z_0}^2 \tag{12-48}$$

and the coefficient of this factor in Eq. (12-47) is called the scattering cross section of the scatterer. It represents the area from which the energy flow of the incident beam is "removed" by the scatterer. The cross section here appears to have the dimensions of length, but this is because we have throughout treated a unit length of the cylinder. The geometrical area of the cylinder per unit length is just $2a$.

In the limit of long wavelengths, for which $ka \to 0$, the expression for the scattering cross section can be simplified by means of the power series expansions for the Bessel functions. The term $m = 0$ and the two terms $m = \pm 1$ contribute equally in this limit, while all other terms involve only higher powers of ka. The result is

$$\sigma = \tfrac{3}{4}\pi^2(ka)^3 a, \tag{12-49}$$

which varies as the inverse cube of the wavelength of the radiation and is smaller than the geometrical cross section of the cylinder. To determine the angular distribution of the scattered radiation we must go back to the time average of the Poynting vector. The leading terms give

$$\sigma(\varphi)\,d\varphi = \frac{\overline{N_r}\,d\varphi}{\overline{N_0}},$$

$$\sigma(\varphi) = \frac{2}{\pi k}\left|\frac{J_0'(ka)}{J_0'(ka) + iN_0'(ka)} - \frac{2J_1'(ka)\cos\varphi}{J_1'(ka) + iN_1'(ka)}\right|^2$$

$$\to \frac{2}{\pi k}\frac{\pi^2(ka)^4}{2^4}(1 - 2\cos\varphi)^2 = \frac{\pi(ka)^3}{8}a(1 - 2\cos\varphi)^2. \tag{12-50}$$

The scattering is thus predominantly backward, with a subsidiary maximum in the forward direction, and zero intensity at $\varphi = 60°$.

12-8 Spherical waves. The scalar wave equation is easily separable in spherical polar coordinates, and the solution of

$$\nabla^2\psi = \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} = -k^2\psi \tag{12-51}$$

is

$$\psi = \frac{1}{\sqrt{kr}}Z_{l+\frac{1}{2}}(kr)Y_l^m(\theta,\varphi)e^{-i\omega t}, \tag{12-52}$$

where $Y^m(\theta,\varphi)$ is, as usual, the spherical harmonic, and $Z_{l+\frac{1}{2}}$ is that Bessel function which is appropriate to the boundary conditions of the given problem. It is convenient to use "spherical Bessel functions" following a terminology* according to which

$$z_l(kr) = \sqrt{\frac{\pi}{2kr}}\, Z_{l+1/2}(kr),$$

so that the solution corresponding to a particular frequency is simply

$$\psi(r,\theta,\varphi,t) = z_l(kr) Y_l^m(\theta,\varphi) e^{-i\omega t}. \qquad (12\text{–}53)$$

But it is not immediately apparent that the scalar solution is at all useful in determining a vector field. In the coordinate systems utilized thus far we have been able to demonstrate simple yet physically interesting results while avoiding the intrinsic difficulty of the vector wave equation very similar to that mentioned in Chapter 8 in connection with the vector Laplace equation: only in Cartesian coordinates, where the coordinate vectors are in the same direction for all points of space, is Laplace's or D'Alembert's equation simply equivalent to three scalar equations for the three components of the vector. Actually, the homogeneous wave equation is completely solvable in a way which makes it easy to impose spherical boundary conditions, but it is necessary to make more explicit a method similar to that used in the cylindrical case. Essentially, it involves expressing a general vector field of zero divergence as a sum of two partial fields.

Consider the function ψ satisfying the scalar equation (12–51). The derivatives of ψ are also solutions, and so is $\mathbf{r} \times \nabla\psi$, as can be seen by direct substitution or by the symmetry properties of the equation as indicated in a problem. If we note that

$$\mathbf{r} \times \nabla\psi = -\nabla \times (\mathbf{r}\psi)$$

we see that the divergence of this vector vanishes, and thus it satisfies a necessary condition on **E** or **H**. If we let

$$\mathbf{E}_1 = \mathbf{r} \times \nabla\psi$$

we may immediately determine the corresponding magnetic field from the relation $i\omega\mathbf{B}_1 = \nabla \times \mathbf{E}_1$. But $\mathbf{E}_1$ cannot represent a general electric field, since it has no component along **r**. A linearly independent solution can be obtained, however, by interchanging the roles of **E** and **H**, i.e.,

$$\mathbf{H}_2 = \mathbf{r} \times \nabla\psi,$$

---

* The spherical Bessel functions $z_l(kr)$ may be specified as $j_l(kr)$ or $n_l(kr)$, just as Z_l is used to represent J_l or N_l. That is, $j_l(kr) = \sqrt{\pi/2kr}\, J_{l+1/2}(kr)$, $n_l(kr) = \sqrt{\pi/2kr}\, N_{l+1/2}(kr)$.

with $\mathbf{E}_2$ derived from the field equation $-i\omega\mathbf{D}_2 = \nabla \times \mathbf{H}_2$. In general, the field is the sum of the two partial fields, and it is evident that boundary conditions at the surface of a sphere are readily imposed. The two types of waves are analogous to the TE and TM waves of the cylindrical case.

The resonant fields of a spherical cavity can be written down immediately in terms of ψ. For the partial field specified by $\mathbf{E}_1$ above z_l must be restricted to j_l in order to ensure regularity at $r = 0$, and must vanish at $r = r_0$, the radius of the cavity, since it is wholly tangential to the perfectly conducting spherical wall. The eigenvalues of k are therefore those for which $j_l(k_{l,n}r_0) = 0$, where $(k_{l,n}r_0)$ is the nth zero of j_l. The allowed frequencies are then given by $\omega^2/c^2 = k_{l,n}^2$. Explicitly, the fields are

$$\begin{aligned}
E_\theta &= -\frac{1}{\sin\theta}\frac{\partial\psi}{\partial\varphi} = \frac{-im}{\sin\theta}\psi,\\
E_\varphi &= \frac{\partial\psi}{\partial\theta},\\
H_r &= \frac{i}{\omega\mu}\frac{l(l+1)}{r}\psi,\\
H_\theta &= \frac{i}{\omega\mu}\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\psi}{\partial\theta}\right),\\
H_\varphi &= \frac{i}{\omega\mu}\frac{1}{r\sin\theta}\frac{\partial}{\partial r}\left(r\frac{\partial\psi}{\partial\varphi}\right) = \frac{-m}{\mu\omega}\frac{1}{r\sin\theta}\frac{\partial}{\partial r}(r\psi).
\end{aligned} \tag{12-54}$$

In H_r the angular differential operator acting on ψ is just that whose eigenvalues are $l(l+1)$, and hence the simplification.

The second set of fields, for which $H_r = 0$, are given formally by Eqs. (12-54) with $\mathbf{E}$ and $\mathbf{H}$ interchanged, except for the change of sign and constants implied in $-i\omega\epsilon\mathbf{E} = \nabla \times \mathbf{H}$. The boundary condition that E_θ and E_φ vanish at $r = r_0$ puts different restrictions on k in this case, however. The equation that must now be satisfied is

$$\left\{\frac{\partial}{\partial r}[rj_l(kr)]\right\}_{r=r_0} = 0,$$

and the allowed frequencies for the two partial fields are correspondingly different.

12-9 Scattering by a sphere. The problem of scattering a plane electromagnetic wave by a sphere can be analyzed in terms of the two independent solutions of the previous section. Since both must be used simultaneously in the representation of the incoming and outgoing waves, it is convenient to define the two functions, which we shall call **M** and **N**, so that they have the same dimensions:

$$\mathbf{M}_{l,m} = \nabla \times (\mathbf{r}\psi_{l,m}), \tag{12-55}$$

$$\mathbf{N}_{l,m} = \frac{\nabla \times \mathbf{M}_{l,m}}{k} = \frac{\nabla \times [\nabla \times (\mathbf{r}\psi_{l,m})]}{k}. \tag{12-56}$$

The reciprocal relation,

$$\mathbf{M}_{l,m} = \frac{\nabla \times \mathbf{N}_{l,m}}{k}, \tag{12-57}$$

follows from the fact that both **M** and **N** satisfy the wave equation, in addition to being functions with vanishing divergence. Note that the $\psi_{l,m}$ appearing in Eqs. (12–55) and (12–56) corresponds to two linearly independent solutions of the scalar wave equation, since the spherical harmonic contains a factor $e^{\pm im\varphi}$. Two vector solutions $\mathbf{M}_{l,m}$ and $\mathbf{N}_{l,m}$ are derivable from each other by a differential operation only if the ambiguity in φ is maintained. The dependence on the azimuthal angle may be written explicitly as sin $m\varphi$ or cos $m\varphi$ to facilitate identification with rectangular components of the fields:

$$\psi_{l,m}\left\{\begin{matrix}\text{even}\\ \text{odd}\end{matrix}\right\} = z_l(kr)P_l^m(\cos\theta)\left\{\begin{matrix}\cos m\varphi\\ \sin m\varphi\end{matrix}\right\} e^{-i\omega t}. \tag{12-58}$$

Then, in spherical component form, the vector solutions are

$$\begin{aligned} e^{i\omega t}\mathbf{M}_{l,m}\left\{\begin{matrix}\text{even}\\ \text{odd}\end{matrix}\right\} = {} & \mp\frac{m}{\sin\theta} z_l(kr)P_l^m(\cos\theta)\left\{\begin{matrix}\sin m\varphi\\ \cos m\varphi\end{matrix}\right\}\hat{\boldsymbol{\theta}} \\ & - z_l(kr)\frac{dP_l^m(\cos\theta)}{d\theta}\left\{\begin{matrix}\cos m\varphi\\ \sin m\varphi\end{matrix}\right\}\hat{\boldsymbol{\varphi}}, \end{aligned} \tag{12-59}$$

$$\begin{aligned} e^{i\omega t}\mathbf{N}_{l,m}\left\{\begin{matrix}\text{even}\\ \text{odd}\end{matrix}\right\} = {} & \frac{l(l+1)}{kr} z_l(kr)P_l^m(\cos\theta)\left\{\begin{matrix}\cos m\varphi\\ \sin m\varphi\end{matrix}\right\}\hat{\mathbf{r}} \\ & + \frac{1}{kr}\frac{d}{dr}[rz_l(kr)]\frac{dP_l^m(\cos\theta)}{d\theta}\left\{\begin{matrix}\cos m\varphi\\ \sin m\varphi\end{matrix}\right\}\hat{\boldsymbol{\theta}} \\ & \mp\frac{m}{kr\sin\theta}\frac{d}{dr}[rz_l(kr)]P_l^m(\cos\theta)\left\{\begin{matrix}\sin m\varphi\\ \cos m\varphi\end{matrix}\right\}\hat{\boldsymbol{\varphi}}, \end{aligned} \tag{12-60}$$

where $\hat{\mathbf{r}}$, $\hat{\boldsymbol{\theta}}$, $\hat{\boldsymbol{\varphi}}$, are the customary unit vectors in the directions of increasing values of the respective coordinates. The functions **M** and **N** are called odd or even according to whether ψ is odd or even in the azimuthal angle φ.

Let us consider a plane-polarized plane wave with the electric vector oriented along x and advancing in the direction of the positive z-axis. The magnetic vector of this incoming wave is then confined to the yz-plane. We must express these fields in terms of the **M**'s and **N**'s. The basic formula for the expansion of a plane wave propagated along the polar axis in spherical coordinates is the scalar series:

$$e^{ikr\cos\theta} = \sum_{l=1}^{\infty} i^l(2l+1)j_l(kr)P_l(\cos\theta). \qquad (12\text{–}61)$$

Now a vector in the x-direction has spherical coordinate components given by

$$\begin{aligned}\mathbf{E}_x &= \mathbf{E}_0 e^{i(kr\cos\theta-\omega t)}\\ &= E_x(\sin\theta\cos\varphi\hat{\mathbf{r}} + \cos\theta\cos\varphi\hat{\boldsymbol{\theta}} - \sin\varphi\hat{\boldsymbol{\varphi}}).\end{aligned} \qquad (12\text{–}62)$$

But

$$\sin\theta e^{ikr\cos\theta} = -\frac{1}{ikr}\frac{\partial}{\partial\theta}(e^{ikr\cos\theta}) \qquad (12\text{–}63)$$

and

$$\frac{dP_l(\cos\theta)}{d\theta} = -P_l^1(\cos\theta). \qquad (12\text{–}64)$$

Equations (12–63) and (12–64) make it possible to identify at once the coefficient of $\mathbf{N}_{l,m}$ in a series expansion of $\mathbf{E}_x$, since **M** has no component along **r**, and we note that the dependence on φ makes it necessary to choose **N**(even). It is possible to determine the coefficient of **M** by consideration of the recurrence relations of the Legendre polynomials and the spherical Bessel functions, but this coefficient may also be written down by inspection if we look at the expansion of the accompanying magnetic field:

$$\begin{aligned}\mathbf{H}_y &= H_y(\sin\theta\sin\varphi\hat{\mathbf{r}} + \cos\theta\sin\varphi\hat{\boldsymbol{\theta}} + \cos\varphi\hat{\boldsymbol{\varphi}})\\ &= -\frac{i}{\mu\omega}\nabla\times\mathbf{E}.\end{aligned} \qquad (12\text{–}65)$$

Again the coefficient of **N** may be identified by examination of N_r with the aid of Eqs. (12–63) and (12–64), and this time the choice of **N**(odd) is necessary. The relation between the electric and magnetic vectors completes the identification, and the final results are

$$\mathbf{E}_x = E_0\sum_{l=1}^{\infty} i^l\frac{2l+1}{l(l+1)}[\mathbf{M}_{l,1}(\text{odd}) - i\mathbf{N}_{l,1}(\text{even})], \qquad (12\text{–}66)$$

$$\mathbf{H}_y = -\frac{kE_0}{\omega\mu}\sum_{l=1}^{\infty} i^l\frac{2l+1}{l(l+1)}[\mathbf{M}_{l,1}(\text{even}) + i\mathbf{N}_{l,1}(\text{odd})]. \qquad (12\text{–}67)$$

As already indicated in Eqs. (12–61), the Bessel function in these expansions must be that which is regular at the origin, namely, $j_l(kr)$. The limitation of m to $m = 1$, evident in the limited dependence of the fields on φ, means that both series are single sums over l.

If this plane wave falls on a sphere of radius a, whose center is at the origin of coordinates, there will be a similar series of functions representing the outgoing wave, with coefficients determined by the boundary conditions at the surface of the sphere. The functions **M** and **N** are again those defined by Eqs. (12–59) and (12–60) with $m = 1$ if the boundary conditions are to be satisfied for all values of the angular variables, but $z_l(kr)$ must be chosen to behave as e^{ikr}/r for large r. This function is

$$h_l^{(1)}(kr) = j_l(kr) + in_l(kr) \simeq (-i)^{l+1}\frac{e^{ikr}}{kr}. \tag{12–68}$$

Formally, we may write the scattered electric field as

$$\mathbf{E}_s = E_0 \sum_{l=1}^{\infty} i^l \frac{2l+1}{l(l+1)} [a_l \mathbf{M}_{l,1}(\text{odd}) - ib_l \mathbf{N}_{l,1}(\text{even})], \tag{12–69}$$

and in view of the reciprocal curl relations,

$$\mathbf{H}_s = -\frac{k}{\omega\mu} E_0 \sum_{l=1}^{\infty} i^l \frac{2l+1}{l(l+1)} [b_l \mathbf{M}_{l,1}(\text{even}) + ia_l \mathbf{N}_{l,1}(\text{odd})]. \tag{12–70}$$

If the sphere is a perfect conductor the sum of the tangential components of the incident and reflected electric fields must vanish at every point on the surface of the sphere for each term of the series. Comparison of Eqs. (12–66) and (12–69) leads to evaluation of the a_l's and b_l's:

$$\begin{aligned} a_l &= -\frac{j_l(ka)}{h_l^{(1)}(ka)}, \\ b_l &= -\left\{\frac{\dfrac{d}{dr}[rj_l(kr)]}{\dfrac{d}{dr}[rh_l^{(1)}(kr)]}\right\}_{r=a}. \end{aligned} \tag{12–71}$$

Before examining these coefficients as functions of the size of the sphere and the wavelength of the radiation, let us look at the behavior of the outgoing fields. From the asymptotic formula for $h_l(kr)$, Eq. (12–68), and the components of **M** and **N**, it is evident that the radially directed fields fall off as $1/r^2$, while the transverse fields behave as $1/r$ for large r. The Poynting vector is thus radial, as would be expected, and the total scattered radiation can be determined by integrating over a sphere of large radius. The evaluation of

$$\frac{dW_s}{dt} = \tfrac{1}{2}\,\mathrm{Re}\int |\mathbf{E}_s \times \mathbf{H}_s^*|\, r^2 \sin\theta\, d\theta\, d\varphi = \tfrac{1}{2}\,\mathrm{Re}\int (E_\theta H_\varphi^* - E_\varphi H_\theta^*) r^2\, d\Omega \tag{12–72}$$

may be accomplished by means of the integral formulas for the associated Legendre functions:

$$\int_0^\pi \left(\frac{dP_l^1}{d\theta}\frac{dP_{l'}^1}{d\theta} + \frac{1}{\sin^2\theta} P_l^1 P_{l'}^1\right)\sin\theta\, d\theta = \frac{2[l(l+1)]^2}{2l+1}\,\delta_{l,l'},$$

$$\int_0^\pi \left(\frac{P_l^1}{\sin\theta}\frac{dP_{l'}^1}{d\theta} + \frac{P_{l'}^1}{\sin\theta}\frac{dP_l^1}{d\theta}\right)\sin\theta\, d\theta = 0.$$

The result of performing the integration indicated in Eq. (12–72) is

$$\frac{dW_s}{dt} = \frac{\pi E_0^2}{k^2}\sqrt{\frac{\epsilon_0}{\mu_0}}\sum_{l=1}^{\infty}(2l+1)(|a_l|^2 + |b_l|^2). \tag{12–73}$$

Since the energy of the incident wave per unit area per unit time is

$$\frac{dW}{dt} = \frac{1}{2} E_0^2 \sqrt{\frac{\epsilon_0}{\mu_0}},$$

the cross section for scattering can be written

$$\sigma = \frac{2\pi}{k^2}\Sigma(2l+1)(|a_l|^2 + |b_l|^2). \tag{12–74}$$

While the coefficients a_l and b_l are tedious to compute, some idea of their general behavior may be gained by writing them in the form

$$a_l = ie^{i\gamma_l}\sin\gamma_l, \quad b_l = ie^{i\gamma_l'}\sin\gamma_l', \tag{12–75}$$

where

$$\tan\gamma_l = \frac{j_l(ka)}{n_l(ka)}, \quad \tan\gamma_l' = \left[\frac{d[rj_l(kr)]}{dr}\bigg/\frac{d[rn_l(kr)]}{dr}\right]_{r=a}.$$

In terms of these parameters, the scattering cross section is

$$\sigma = \frac{2\pi}{k^2}\sum_l (2l+1)(\sin^2\gamma_l + \sin^2\gamma_l'), \tag{12–76}$$

from which it is clear that the contributions of the two kinds of waves whose electric fields are represented by **M** and **N** oscillate as functions of ka. The total cross section reduces to a very simple expression only when $ka \ll 1$. In that case, the first term in the series expansions for the Bessel functions gives

$$|a| = \frac{(ka)^3}{3}, \quad |b| = \frac{2(ka)^3}{3},$$

so that

$$\sigma = \frac{10\pi}{3} k^4 a^6. \tag{12–77}$$

Thus, in the limit of low frequencies, the scattering cross section varies as the inverse fourth power of the wavelength, and is, as in the corresponding cylindrical scattering case, smaller than the geometrical cross section.

The scattering of a plane wave by a sphere of general electric and magnetic properties embedded in a dielectric follows the same pattern as the solution outlined here. The fields inside the sphere must also be considered, but the condition that the tangential components of both **E** and **H** be continuous across the boundary is sufficient to determine all the coefficients. Again in the low-frequency limit it is found that the cross section varies as the inverse fourth power of the wavelength.

SUGGESTED REFERENCES

J. A. Stratton, *Electromagnetic Theory.* The literature on waves in regions with metallic boundaries has grown very rapidly, but the principles and many of the methods involved in applications are treated in Chapter IX of this work. Chapters VI and VII deal with cylindrical and spherical waves.

P. M. Morse and H. Feshbach, *Methods of Theoretical Physics.* This comprehensive and elegant presentation of mathematical methods contains all one might hope to know on boundary value problems. The authors regret that it was not available during the writing of the present work!

E. U. Condon, Principles of Microwave Radio, *Rev. Modern Phys.* **14** (1942). An early review article, forming an excellent introduction to the physics of cavity and guide modes.

J. C. Slater, *Microwave Electronics.* An expanded and more recent (1950) version of a *Rev. Modern Phys.* (1946) article by the same author, stressing physical principles.

P. M. Morse, *Vibration and Sound,* and H. Lamb, *Hydrodynamics* are interesting for comparing the cylindrical scattering of longitudinal waves with the results of Section 12–7.

M. Born, *Optik.* Chapter VI contains a full account of spherical scattering, including the angular distribution for various ratios of wavelength to diameter.

A. Sommerfeld, *Electrodynamics.* The boundary value problem emphasized here is that of waves on wires, and thus the work supplements both the present chapter and the references above.

E. L. Ince, *Ordinary Differential Equations.* A discussion of the eigenvalue problem will be found in Chapter 10 of this work.

E. Jahnke and F. Emde, *Tables of Functions.* For numerical applications this is the best and most convenient short table of functions.

G. N. Watson, *Theory of Bessel Functions.* The properties of Bessel functions will be found in many references, but for the proofs for series expansions and similar matters Watson is excellent.

EXERCISES

1. Show that at a conducting boundary the ratio of the normal component of the magnetic field to its tangential component is of the order of the skin depth to the wavelength of the oscillating fields.

2. A rectangular cavity of dimensions a, a, L, and walls of large conductivity σ is excited in the mode

$$E_z = E_0 \sin(\pi x/a) \sin(\pi y/a) e^{-i\omega t},$$

$$H_z = E_x = E_y = 0.$$

What is ω? Calculate the forces exerted on all the faces. What is Q for this mode?

3. Show that the absolute cutoff for TE modes in a wave guide of rectangular cross section is lower than for TM modes. Find the attenuation coefficient for the lowest TE mode.

4. Find the lowest frequency for a cylindrical cavity of circular cross section in terms of its radius and height. How would you estimate Q for this mode?

5. Find the fields and cross section for scattering from a conducting cylinder in the limit $ka \ll 1$ if the incoming wave is polarized so that E is along z, parallel to the axis of the cylinder.

6. A solution of the wave equation for a given orientation of the coordinate axes must be a solution for any other orientation of the axes. Consider the difference between $\psi(r,\theta,\varphi)$ and $\psi(r,\theta',\varphi')$ for an infinitesimal rotation to show that $\mathbf{r} \times \nabla\psi$ is a solution if ψ is a solution.

7. Calculate the total force exerted on the reflecting sphere of Section 12–8 under the conditions that $ka \ll 1$. Estimate the change in your result if the sphere were of copper instead of being a perfect conductor.

8. Consider a cubical cavity with perfectly conducting walls. Prove that for waves short in comparison with the dimension l of the cube the number of modes within a frequency range $d\omega$ is given by

$$\frac{l^3\omega^2\, d\omega}{\pi^2 c^3},$$

or for $d\nu$ by

$$\frac{l^3 8\pi\nu^2\, d\nu}{c^3}.$$

Does the number of modes per unit volume depend on the shape of the cavity? This formula is of great historical importance in the theory of "blackbody" radiation.

CHAPTER 13

THE INHOMOGENEOUS WAVE EQUATION

13–1 The wave equation for the potentials. The relation of radiation fields to their sources is most readily found in terms of potential functions. In view of the law of induction, curl **E** is not zero in a changing magnetic field, so that it is no longer possible to derive the electric field solely from a scalar potential. The divergence of **B** vanishes under all conditions, however, and thus the magnetic field is still derivable from a vector potential:

$$\mathbf{B} = \nabla \times \mathbf{A}. \tag{7–37}$$

If Eq. (7–37) is assumed valid, then the electric field can be written as the sum of the gradient of a scalar potential and a supplementary nonconservative contribution from the rate of change of the vector potential. That is, **E** may be derived from the scalar and vector potentials by

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}, \tag{13–1}$$

which makes **E** conform to the relation of the third of Eqs. (9–7). As before, the divergence of **A** remains undefined, or at least remains undefined within an additive arbitrary function of position. Let us define the divergence of **A** by what is called the Lorentz condition,

$$\nabla \cdot \mathbf{A} + \mu\epsilon \frac{\partial \phi}{\partial t} + \mu\sigma\phi = 0, \tag{13–2}$$

which in free space becomes

$$\nabla \cdot \mathbf{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t} = 0. \tag{13–3}$$

The Lorentz condition is not quite the arbitrary subsidiary relation it appears at first sight. We shall see at once that it has the advantage of introducing complete symmetry between the scalar and vector potentials, i.e., it makes both potentials satisfy the same wave equation as that obeyed by the fields. Later we shall find that the Lorentz condition also assures a relativistic covariant relation between the scalar and vector potentials. To determine the equations satisfied by the potentials we need only introduce the defining equations (7–37) and (13–1) and the Lorentz condition into the first and fourth of Maxwell's equations, Eqs. (9–7).

We obtain the symmetrical set of equations:

$$\nabla^2\mathbf{A} - \mu\epsilon\frac{\partial^2\mathbf{A}}{\partial t^2} - \mu\sigma\frac{\partial\mathbf{A}}{\partial t} = -\mu\mathbf{j}', \tag{13-4}$$

$$\nabla^2\phi - \mu\epsilon\frac{\partial^2\phi}{\partial t^2} - \mu\sigma\frac{\partial\phi}{\partial t} = -\rho/\epsilon. \tag{13-5}$$

Here $\mathbf{j}'$ represents a current given by $\mathbf{j}' = \sigma\mathbf{E}'$, that part of the current density that is produced by the external electromotive forces. It does not contain any part of the current induced by the electric fields in conducting media and represented by the last term on the left side of Eq. (13-4).

Let us introduce the symbolic operator $\Box$, known as the D'Alembertian, and defined by

$$\Box = \nabla^2 - \mu\epsilon\frac{\partial^2}{\partial t^2}. \tag{13-6}$$

In free space, Eqs. (13-4) and (13-5) then become simply

$$\Box\mathbf{A} = -\mu_0\mathbf{j}', \tag{13-7}$$

$$\Box\phi = -\rho/\epsilon_0, \tag{13-8}$$

where $\mathbf{j}'$ and ρ are the sources of the field and are produced by external agents. These equations are known as the inhomogeneous wave equations. Their particular solutions are expressible in terms of integrals over the charge and current distributions, while their complementary solutions, namely, those of the homogeneous equations, are obviously just the wave solutions.

By inspection of Eqs. (7-37) and (13-1) it can be seen that the resultant electric and magnetic fields are unchanged by transformations of the type

$$\mathbf{A}' = \mathbf{A} - \nabla\psi, \tag{13-9}$$

$$\phi' = \phi + \frac{\partial\psi}{\partial t}, \tag{13-10}$$

where ψ is a function of the coordinates and the time. This means that if any physical law involving electromagnetic interaction is to be expressed in terms of the general electrodynamic potentials $\mathbf{A}$ and ϕ then such a physical law must be unaffected by a transformation of the type given by Eqs. (13-9) and (13-10). These transformations are usually known as gauge transformations, and a physical law that is invariant under such a transformation is said to be gauge invariant. The property of gauge invariance ensures that the physical law will not lead to consequences that cannot be expressed in the field formulation of the interaction of charges and currents.

13–2 Solution by Fourier analysis. Let us begin the investigation of the particular solutions of the inhomogeneous equations by reviewing the solution of the analogous static problem. In the static case, Eq. (13–8) reduces to Poisson's equation,

$$\nabla^2\phi = -\rho/\epsilon_0, \tag{1–35}$$

of which the particular solution is

$$\phi(x_\alpha) = \frac{1}{4\pi\epsilon_0}\int \frac{\rho(x'_\alpha)}{r(x_\alpha,x'_\alpha)}\,dv'.$$

What we seek now is the modification of the solution of Eq. (1–35) that is produced by the presence of the time-dependent term in Eq. (13–7).

The equations for both ϕ and $\mathbf{A}$ have the general form

$$\Box\psi(x_\alpha,t) = -g(x_\alpha,t). \tag{13–11}$$

Let us assume that the source function $g(x_\alpha,t)$ can be analyzed by the Fourier integral

$$g(x_\alpha,t) = \int_{-\infty}^{\infty} g_\omega(x_\alpha)e^{-i\omega t}\,d\omega, \tag{13–12}$$

which has the Fourier inversion

$$g_\omega(x_\alpha) = \frac{1}{2\pi}\int_{-\infty}^{\infty} g(x_\alpha,t)e^{i\omega t}\,dt. \tag{13–13}$$

Similarly, we may analyze the general potential $\psi(x_\alpha,t)$ into Fourier components,

$$\psi(x_\alpha,t) = \int_{-\infty}^{\infty} \psi_\omega(x_\alpha)e^{-i\omega t}\,d\omega, \tag{13–14}$$

with a corresponding inverse relation

$$\psi_\omega(x_\alpha) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \psi(x_\alpha,t)e^{i\omega t}\,dt. \tag{13–15}$$

By substitution of Eqs. (13–12) and (13–14) into Eq. (13–11), we see that the Fourier component $\psi_\omega(x_\alpha)$ obeys the differential relation

$$\nabla^2\psi_\omega + \frac{\omega^2}{c^2}\psi_\omega = -g_\omega, \tag{13–16}$$

which is similar to Poisson's equation.

We may synthesize the solution of Eq. (13–16) by a superposition of unit point solutions corresponding to a source at the point x'_α given by

$g_\omega(x_\alpha) = \delta(x_\alpha - x'_\alpha)$, where $\delta(x_\alpha - x'_\alpha)$ is the Dirac δ-function. Each unit source potential satisfies the equation

$$\nabla^2 G(x_\alpha, x'_\alpha) + \frac{\omega^2}{c^2} G(x_\alpha, x'_\alpha) = -\delta(x_\alpha - x'_\alpha), \tag{13-17}$$

with G a function of both x_α and x'_α. The partial solution corresponding to the frequency ω of the total source is then given by the superposition

$$\psi_\omega(x_\alpha) = \int g_\omega(x'_\alpha) G(x_\alpha, x'_\alpha)\, dv'. \tag{13-18}$$

To find G, we note that the solution of Eq. (13-17) will be spherically symmetric in r, the distance between points x_α and x'_α, and hence at all points other than $r = 0$ it will be identical with the solution of the equation

$$\frac{1}{r}\frac{d^2(rG)}{dr^2} + k^2 G = 0,$$

where $k = \omega/c$ as usual. This equation can be integrated immediately:

$$G = \frac{A}{r} e^{\pm ikr}.$$

To evaluate the constant A, we consider the volume integral of Eq. (13-17) over the neighborhood of the singular point $r = 0$. Here G behaves as A/r, and the use of Eq. (1-11) in the integration leads to $-4\pi A = -1$. Therefore

$$G = \frac{1}{4\pi r} e^{\pm ikr} \tag{13-19}$$

is the solution of Eq. (13-17), and substitution into Eq. (13-18) gives

$$\psi_\omega(x_\alpha) = \frac{1}{4\pi} \int \frac{g(x'_\alpha)}{r(x_\alpha, x'_\alpha)} e^{\pm ikr(x_\alpha, x'_\alpha)}\, dv'. \tag{13-20}$$

The effect of the second term in Eq. (13-16) is then simply to introduce a complex exponential factor in the Coulomb potential expression of Chapter 1.

In terms of Eq. (13-20) the time-dependent potential satisfying Eq. (13-11) is given by

$$\psi(x_\alpha, t) = \int \psi_\omega(x_\alpha) e^{-i\omega t}\, d\omega = \frac{1}{4\pi} \iint \frac{g_\omega(x'_\alpha) e^{-i(\omega t \pm kr)}}{r(x_\alpha, x'_\alpha)}\, d\omega\, dv'. \tag{13-21}$$

We may introduce a new time

$$t'(x_\alpha, x'_\alpha) = t \pm r/c = t \pm kr/\omega, \tag{13-22}$$

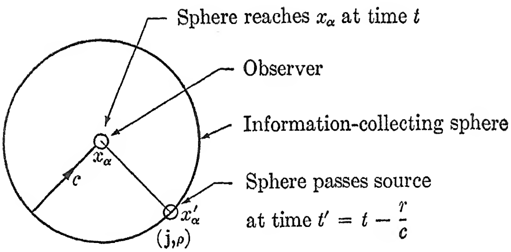

Sphere reaches x_α at time t
Observer
Information-collecting sphere
x_α
c
Sphere passes source at time $t' = t - \frac{r}{c}$
x'_α
(j,ρ)

FIG. 13–1 Sphere collecting information for the retarded potential at x_α.

which corresponds to shifting the origin of time by an amount equal to the time it takes a light signal to be propagated from point x'_α to point x_α. If we now evaluate the Fourier transformation by using Eq. (13–12), we obtain

$$\psi(x_\alpha,t) = \frac{1}{4\pi}\int \frac{g(x'_\alpha,t \pm r/c)}{r(x_\alpha,x'_\alpha)}\,dv'. \tag{13–23}$$

Note that the point of observation is contained explicitly in the denominator of the integrand and also, by means of Eq. (13–22), in the time at which the time-varying currents and charges are introduced into the integration. Mathematically, both plus and minus signs in Eq. (13–23) are valid, but only the minus sign appears to have physical significance. We are concerned with the effect at x_α of sources at x'_α, and the minus sign corresponds to the cause preceding the effect. Equation (13–23) with the minus sign is known as the retarded potential solution of the inhomogeneous wave equation. The solution with the plus sign is known as the advanced potential and appears to have no physical significance, although at various times attempts have been made to use the advanced potential to solve certain difficulties in electrodynamics.

A retarded potential might be visualized as follows. Consider an observer located at the point x_α in space as seen in Fig. 13–1, and let a sphere whose center is at x_α contract with a radial velocity c such that it has just converged on the point at the time of observation t. The time at which this information-collecting sphere passes the source of the electric field at point x'_α is then the time at which the source produced the effect which is felt at x_α at time t.

If we denote by the rectangular bracket symbol [ ] that the variables contained within the bracket are to be evaluated at the retarded time t', then the integrals of the inhomogeneous wave equations (13–7) and (13–8) corresponding to a current distribution $\mathbf{j}(x'_\alpha)$ and a charge distribution $\rho(x'_\alpha)$ may be written

$$\mathbf{A}(x_\alpha,t) = \frac{\mu_0}{4\pi}\int \frac{[\mathbf{j}(x'_\alpha)]}{r(x_\alpha,x'_\alpha)}\,dv', \tag{13–24}$$

$$\phi(x_\alpha,t) = \frac{1}{4\pi\epsilon_0}\int \frac{[\rho(x'_\alpha)]}{r(x_\alpha,x'_\alpha)}\,dv'. \tag{13-25}$$

13-3 Fourier components of the fields and radiation.

If the time variation of the scalar and vector potentials is assumed to be

$$\mathbf{A}(x_\alpha,t) = \mathbf{A}_\omega(x_\alpha)e^{-i\omega t}, \quad \phi(x_\alpha,t) = \phi_\omega(x_\alpha)e^{-i\omega t},$$

then the coordinate dependence of a single sinusoidal component of the solutions above, Eqs. (13–24) and (13–25), is given by

$$\mathbf{A}_\omega(x_\alpha) = \frac{\mu_0}{4\pi}\int \frac{\mathbf{j}_\omega e^{ikr}}{r}\,dv', \tag{13-26}$$

$$\phi_\omega(x_\alpha) = \frac{1}{4\pi\epsilon_0}\int \frac{\rho_\omega e^{ikr}}{r}\,dv'. \tag{13-27}$$

The electromagnetic fields are derived from the potentials by Eqs. (7–37) and (13–1). The gradient of the "monochromatic" scalar potential of Eq. (13–27), after we have commuted the ∇ operator with ρ, and changed it to ∇' to operate on the coordinates of the source point, becomes

$$-\nabla\phi_\omega(x_\alpha) = \frac{1}{4\pi\epsilon_0}\int \rho_\omega\nabla'\left(\frac{e^{ikr}}{r}\right)dv'. \tag{13-28}$$

The electric field is therefore

$$\mathbf{E}_\omega(x_\alpha) = \frac{1}{4\pi\epsilon_0}\int \rho_\omega\nabla'\left(\frac{e^{ikr}}{r}\right)dv' + \frac{i\mu_0\omega}{4\pi}\int \mathbf{j}_\omega\frac{e^{ikr}}{r}\,dv'. \tag{13-29}$$

In a similar way, the magnetic field is found to be

$$\mathbf{B}_\omega = \frac{\mu_0}{4\pi}\int \mathbf{j}_\omega \times \nabla'\left(\frac{e^{ikr}}{r}\right)dv'. \tag{13-30}$$

Just as in the analogous static cases, it has been possible here to simplify the expressions by changing from ∇ to ∇' after writing the vector operator so that it does not operate on the source densities.

We see that under the modification of retardation the integrals involve the gradient operator acting on e^{ikr}/r instead of simply $1/r$. This means that when the differentiation is carried out two terms will result, one varying as $1/r$ and the other as $1/r^2$. We have already seen that the $1/r$ term might give a finite contribution in terms of the surface integral of the Poynting vector, and that it therefore represents a net radiation energy loss. The region in which the $1/r$ term is dominant is known as the wave zone, or radiation field zone, while the zones of higher inverse order radial dependence are known as the quasi-stationary or inductance field zones.

From the fields of Eqs. (13–29) and (13–30) we may compute the total energy loss by radiation for an arbitrary current distribution $\mathbf{j}(x'_\alpha)$. If only the $1/r$ dependent radiation field is taken into account, the magnetic field may be written

$$\mu_0 \mathbf{H}_\omega = \mathbf{B}_\omega = -\frac{i\mu_0}{4\pi}\int (\mathbf{j}_\omega \times \mathbf{k})\frac{e^{ikr}}{r}\,dv', \tag{13–31}$$

since the propagation vector $\mathbf{k}$ is in the same direction as $\mathbf{r}$. By comparing the current-dependent part of Eq. (13–29) with Eq. (13–31) it is seen that in the radiation zone $\mathbf{E}$ and $\mathbf{H}$ are at right angles to each other, and that the ratio of the magnitudes is given by $R_0 = \sqrt{\mu_0/\epsilon_0}$ [cf. Eq. (11–10)]. The time average of the Poynting vector is therefore

$$\overline{\mathbf{N}} = \tfrac{1}{2}\,\mathrm{Re}\,(\mathbf{E} \times \mathbf{H}^*) = \tfrac{1}{2}R_0|\mathbf{H}|^2\hat{\mathbf{r}}$$

where $\hat{\mathbf{r}} = \mathbf{r}/r$. With Eq. (13–31) for $\mathbf{H}_\omega$, we obtain

$$\overline{\mathbf{N}_\omega} = \frac{R_0}{32\pi^2}\left|\int (\mathbf{j}_\omega \times \mathbf{k})\frac{e^{ikr}}{r}\,dv'\right|^2\hat{\mathbf{r}} \tag{13–32}$$

as a function of the position of the observation point.

In general, this formula does permit the calculation of the rate of radiation from a given system of currents. If the point of observation is at a distance r large compared with the dimension of the radiating system, it is convenient to refer all coordinates to an origin within the current distribution, as in Fig. 13–2. Then, since $\mathbf{r} = \mathbf{R} - \boldsymbol{\xi}$, Eq. (13–31) becomes

$$\mathbf{B}_\omega = -\frac{i\mu_0}{4\pi}\frac{e^{ikR}}{R}\int (\mathbf{j}_\omega \times \mathbf{k})e^{-i\mathbf{k}\cdot\boldsymbol{\xi}}\,dv', \tag{13–33}$$

since ξ is small compared with R. The rate of total energy radiation can then be expressed as an integral over a sphere of radius R which reduces

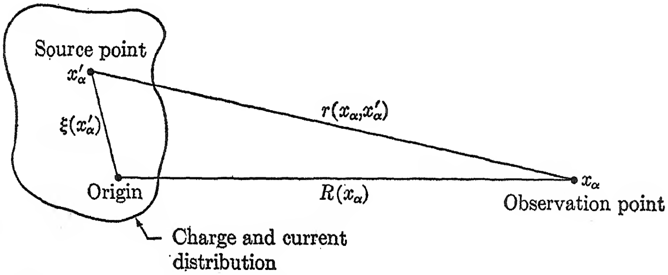


FIG. 13–2 Source and field coordinates referred to an origin within the charge and current distribution.

to an integral over the solid angle Ω subtended at the source. The final relation for the rate of total radiated energy corresponding to a system radiating at a single frequency ω is

$$U_\omega = \int \mathbf{N}_\omega \cdot d\mathbf{S} = \frac{R_0}{32\pi^2} \int \left| \int (\mathbf{j}_\omega \times \mathbf{k}) e^{-i\mathbf{k}\cdot\boldsymbol{\xi}}\, dv' \right|^2 d\Omega, \qquad (13\text{–}34)$$

where $d\Omega = \mathbf{R} \cdot d\mathbf{S}/R^3$.

Let us apply this formula to a case where the dimensions of the radiating system are also small compared with the wavelength, so that the exponential factor in the integrand is approximately 1. In particular, consider a linear current distribution given by

$$J = J_0 \cos \frac{2\pi\xi}{\lambda}$$

oscillating at a single frequency corresponding to the free space wavelength λ. If θ is the angle between the line current and $\mathbf{k}$, the volume integral over the current distribution of length L reduces to

$$kJ_0 \sin\theta \int_{-L/2}^{L/2} \cos\left(\frac{2\pi\xi}{\lambda}\right) d\xi = kJ_0 L \frac{\sin(\pi L/\lambda)}{\pi L/\lambda} \sin\theta.$$

Therefore

$$U_\omega = \frac{\pi}{6} R_0 \left(\frac{J_0 L}{\lambda}\right)^2 \left[\frac{\sin(\pi L/\lambda)}{\pi L/\lambda}\right]^2, \qquad (13\text{–}35)$$

which for L small in comparison with λ reduces to

$$U_\omega = \frac{\pi}{6} R_0 \left(\frac{L}{\lambda}\right)^2 J_0^2.$$

With an ohmic resistance R the same alternating current would have a heat loss $RJ_0^2/2$; by analogy it is customary to define "radiation resistance" by

$$U_\omega = R_{\text{rad}} J_0^2/2,$$

and we see that for a short radiator

$$R_{\text{rad}} = \frac{\pi}{3} R_0 \left(\frac{L}{\lambda}\right)^2.$$

In general, Eq. (13–34) refers only to a nonchromatic source. To find the total radiation from a pulse of current we must use $\mathbf{H}$ instead of $\mathbf{H}_\omega$

and integrate over the time.*

$$\int_{-\infty}^{\infty} H^2(t)\,dt = \int_{-\infty}^{\infty} dt \left(\int_{-\infty}^{\infty} H_\omega e^{-i\omega t}\,d\omega\right)\left(\int_{-\infty}^{\infty} H_{\omega'} e^{-i\omega' t}\,d\omega'\right)$$

$$= \int_{-\infty}^{\infty} H_\omega\,d\omega \int_{-\infty}^{\infty} H_{\omega'}\,d\omega' \int_{-\infty}^{\infty} e^{-i(\omega+\omega')t}\,dt$$

$$= 2\pi \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H_\omega H_{\omega'}\,\delta(\omega+\omega')\,d\omega\,d\omega'$$

$$= 2\pi \int_{-\infty}^{\infty} H_\omega H_{-\omega}\,d\omega = 4\pi \int_0^{\infty} |H_\omega|^2\,d\omega, \tag{13-36}$$

since $H_{-\omega} = H_\omega^*$. Hence the energy loss corresponding to a given frequency interval $d\omega$ is $4\pi U_\omega\,d\omega$, or the total energy loss during the pulse is

$$\delta W = \frac{R_0}{8\pi} \iint \left| \int (\mathbf{j}_\omega \times \mathbf{k}) e^{-i\mathbf{k}\cdot\boldsymbol{\xi}}\,dv' \right|^2 d\Omega\,d\omega. \tag{13-37}$$

The general case in which the dimensions of the source are small compared with λ, so that the factor $e^{-i\mathbf{k}\cdot\boldsymbol{\xi}}$ can be expanded in a power series about the center of the system, will be considered in the following sections. It is clear that the various terms in such a power series expansion will represent energy losses that depend on successively higher powers of the frequency.

In this discussion we have omitted the induction field altogether, although we have not proved explicitly that it does not contribute to the radiated energy. What we have proved is that the induction field cannot contribute to the energy radiated over a surface of very large radius, since its higher order of inverse radial dependence causes the surface integrals of the induction field to vanish. The evaluation of the Poynting vector of the induction fields over a finite surface would give a nonzero instantaneous energy flow, although the time average of each Fourier component vanishes. This implies that the induction fields give rise to energy fluctuations in the radiation field. In particular, if we have a radiating system whose oscillations are nonsinusoidal, so that it undergoes a net change in its configuration, then net energy can be transferred into the induction field.

---

* We are here evaluating $\int_{-\infty}^{\infty} H^2(t)\,dt$ for a *real* field $\mathbf{H}(t) = \int_{-\infty}^{\infty} \mathbf{H}_\omega e^{-i\omega t}\,d\omega$. This implies that in general $\mathbf{H}_\omega$ is complex. On the other hand, in treating quantities of purely sinusoidal time variation [such as in Eqs. (13-26) to (13-34)], an expression of the type $\mathbf{H}_\omega e^{-i\omega t}$ actually denotes its real part.

13–4 The Hertz potential. The Lorentz condition of Eq. (13–3) which relates the scalar and vector potentials is consistent with the equation of continuity,

$$\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0. \tag{7–1}$$

In fact, it can be easily shown that if the retarded potentials of Eqs. (13–24) and (13–25) are assumed, then the Lorentz condition is a direct consequence of the equation of continuity. Since charges and currents cannot be specified independently, it is advantageous for the calculation of radiation fields to represent them by a single function chosen in such a way that the continuity equation is identically satisfied. The radiation field can similarly be represented by a single potential so chosen that the Lorentz condition is identically satisfied. The first requirement can be met by deriving the charge and current densities from a single vector function $\mathbf{p}(x'_\alpha,t)$ of the source coordinates by the relations

$$\rho_{\text{true}} = -\nabla \cdot \mathbf{p}, \tag{13–38}$$

$$\mathbf{j}_{\text{true}} = \frac{\partial \mathbf{p}}{\partial t}. \tag{13–39}$$

It is seen by inspection that the equation of continuity is satisfied by this choice of **p**, but otherwise an arbitrary current distribution can be represented. **p** is known as the polarization vector, and it is related to the true charges and true currents in the same way that the dielectric polarization **P** is related to the polarization charges and polarization currents. This is only a mathematical parallel, however, and it should be emphasized that **j** and ρ represent the true charges which constitute the external sources of the field, so that **p** is entirely different from the ordinary polarization **P**.

A vector $\mathbf{\Pi}$ which combines in a similar way the potentials **A** and ϕ, and which at the same time implies the Lorentz condition, is defined by the equations

$$\mathbf{A} = \frac{1}{c^2}\frac{\partial \mathbf{\Pi}}{\partial t}, \quad \phi = -\nabla \cdot \mathbf{\Pi}. \tag{13–40}$$

$\mathbf{\Pi}$ is known as the polarization potential, or Hertz vector. Since the operation indicated by Eqs. (13–40) is linear, $\mathbf{\Pi}$ will obey the homogeneous wave equation in source-free space, $\Box \mathbf{\Pi} = 0$. By combining the definitions of **p** and $\mathbf{\Pi}$, we find that if the Hertz vector obeys the inhomogeneous wave equation with **p** as the source, then the ordinary potentials ϕ and **A** obey their respective wave equations with ρ and **j** as sources. That is, we have

$$\Box \mathbf{\Pi} = \nabla^2 \mathbf{\Pi} - \frac{1}{c^2}\frac{\partial^2 \mathbf{\Pi}}{\partial t^2} = -\frac{\mathbf{p}}{\epsilon_0}. \tag{13–41}$$

Note that this is a three-component equation, whereas **A** and ϕ amount to a four-component potential.

The retarded potential solution of Eq. (13–41) is given by

$$\mathbf{\Pi}(x_\alpha) = \frac{1}{4\pi\epsilon_0}\int \frac{[\mathbf{p}(x'_\alpha)]}{r(x_\alpha,x'_\alpha)}\,dv' \tag{13–42}$$

for arbitrary time variation, and by

$$\mathbf{\Pi}_\omega = \frac{1}{4\pi\epsilon_0}\int \frac{\mathbf{p}(x'_\alpha)e^{ikr}}{r(x_\alpha,x'_\alpha)}\,dv' \tag{13–43}$$

for sinusoidal time variation. The fields can be derived from the Hertz vector by means of the defining equations, Eqs. (13–40). If we let

$$\mathbf{C} = \nabla \times \mathbf{\Pi}, \tag{13–44}$$

the magnetic field is given by

$$\mathbf{B} = \frac{1}{c^2}\frac{\partial \mathbf{C}}{\partial t} \tag{13–45}$$

and the electric field by

$$\mathbf{E} = \nabla \times \mathbf{C}. \tag{13–46}$$

Equation (13–45) is quite general, but Eq. (13–46) is true only outside the source, where $\nabla \cdot \mathbf{E} = 0$.

13–5 Computation of radiation fields by the Hertz method. We shall apply this method to the case where the fields are observed at distances large compared with the extent of the source distribution, and where in addition the extent of the source is reasonably small compared with the wavelength of the outgoing radiation. These restrictions, $\xi(x'_\alpha) \ll r(x_\alpha,x'_\alpha)$, $\xi(x'_\alpha) \ll \lambda\!\!\!^{-}$, are equivalent to assuming that the phase shift of the outgoing wave over the current and charge distributions is small, and that the distance to the observer is large compared to the dimensions of the source.

Under these conditions, the function e^{ikr}/r is a slowly varying function relative to the variation of **p** itself in the integral of Eq. (13–43). It is therefore natural to consider the expansion of this function as a power series in terms of the distance ξ of the source point from the origin of the charge distribution, and to study the asymptotic behavior of this expansion for large distances R of the observation point. The parameters of expansion will be taken as ratios of ξ and R to the wavelength $\lambda\!\!\!^{-} = 1/k$. Such an expansion can be derived by considering the expansion of e^{ikr}/r relative to a shifted origin:

$$\frac{e^{ikr}}{r} = ik\sum_{n=0}^{\infty}(2n+1)P_n(\cos\theta)j_n(k\xi)h_n^{(1)}(kR). \tag{13–47}$$

In accord with our approximations, we are interested in the behavior of Eq. (13–47) for small values of ξ and large values of R. We have previously noted that the asymptotic behavior of the R-dependent function for $kR \gg 1$ is

$$h_n^{(1)}(kR) \simeq (-i)^{n+1} \frac{e^{ikR}}{kR}. \tag{12–68}$$

Thus we obtain for the vector $\mathbf{\Pi}_\omega$ of Eq. (13–43),

$$\mathbf{\Pi}_\omega(x_\alpha) = \frac{e^{ikR}}{4\pi\epsilon_0 R} \int \sum_{n=0}^{\infty} (-i)^n (2n+1) P_n(\cos\theta) j_n(k\xi) \mathbf{p}_\omega \, dv'. \tag{13–48}$$

Since the Legendre functions are orthogonal, there will be no interference between terms of the series when the field quantities are squared and integrated over solid angle, and we may neglect the phase shifts. The relative magnitude of successive terms will depend on the symmetry properties of the charge distribution, but from the fact that for $k\xi \ll 1$

$$j_n(k\xi) \simeq \frac{2^n n!}{(2n+1)!} (k\xi)^n \tag{13–49}$$

it is evident that the lowest nonvanishing term is the only one of importance if the wavelength is long compared with the dimensions of the source. Under these conditions it is customary to use the approximation of Eq. (13–49) and retain only the largest term:

$$\mathbf{\Pi}_\omega^{(n)} = \frac{e^{ikR}}{4\pi\epsilon_0 R} \frac{2^n n!}{(2n)!} \int (k\xi)^n \mathbf{p}_\omega P_n(\cos\theta) \, dv'. \tag{13–50}$$

This formulation expresses the radiation field as a moment of the polarization vector. If the source distribution is linear in space the Legendre function of the angle between the field point and source point at the arbitrary origin may be taken outside the integral sign, so that it automatically defines the angular distribution of the corresponding fields. Otherwise the addition theorem for Legendre functions must be used to ascertain the dependence of the radiation field on its angular coordinates, but the lowest terms are easy to handle. In any case the zeroth term, $n = 0$, involves just $\mathbf{p}_1 = \int \mathbf{p}_\omega \, dv'$, the same total dipole moment of the distribution as that discussed in Chapter 1. The next higher terms will become important only if the dipole term vanishes, that is, only if the distribution does not have a net oscillating dipole moment. Note, however, that if more than one moment contributes to the radiation we *cannot* simply add the corresponding contributions as given by Eq. (13–50); the approximation (13–49) is not valid for any but the highest contributing moment, since the higher terms of its expansion will have the same effect as a contribution of terms of larger n.*

---

* This effect is illustrated in Problem 8, following.

The multipole expansion can also be arrived at by a purely formal Taylor expansion in powers of the source coordinates ξ. If $\xi \ll \lambda$ within the region where $\mathbf{p}$ is different from zero, we may expand Eq. (13–43) to obtain

$$\begin{aligned} 4\pi\epsilon_0 \Pi_{\omega\beta}(x_\alpha) &= \int p_\beta(x'_\alpha)\,\frac{e^{ikr}}{r}\,dv' \\ &= \frac{e^{ikR}}{R}\int p_\beta\,dv' - \left[\frac{\partial}{\partial x_\gamma}\left(\frac{e^{ikr}}{r}\right)\right]_{r=R}\int \xi_\gamma p_\beta\,dv \\ &\quad + \frac{1}{2!}\left[\frac{\partial}{\partial x_\gamma}\frac{\partial}{\partial x_\delta}\left(\frac{e^{ikr}}{r}\right)\right]_{r=R}\int \xi_\gamma\xi_\delta p_\beta\,dv' \cdots . \end{aligned}$$

The first two terms will be readily identified with Eqs. (13–51) and (13–63) below as calculated from Eq. (13–50).

13–6 Electric dipole radiation. Let us consider the dipole field in detail. For $n = 0$ the Hertz vector, $\mathbf{\Pi}_\omega$, is given by

$$\mathbf{\Pi}_\omega(x_\alpha) = \frac{e^{ikR}}{4\pi\epsilon_0 R}\int \mathbf{p}_\omega(x'_\alpha)\,dv' = \frac{e^{ikR}}{4\pi\epsilon_0 R}\,\mathbf{p}_1, \tag{13–51}$$

and we may take the direction of $\mathbf{p}_1$ as parallel to the polar axis. To obtain the radiation field from the polarization potential of Eq. (13–51), we shall first find $\mathbf{C}$ as defined by Eq. (13–44). From the components of the polarization vector as shown in Fig. 13–3, it is clear that the spherical polar components of $\mathbf{\Pi}_{0\omega}$ are

$$\begin{aligned} \Pi_R &= \Pi_{0\omega}\cos\theta = \frac{p_1\cos\theta e^{ikR}}{4\pi\epsilon_0 R}, \\ \Pi_\theta &= -\Pi_{0\omega}\sin\theta = -\frac{p_1\sin\theta e^{ikR}}{4\pi\epsilon_0 R}, \\ \Pi_\varphi &= 0. \end{aligned} \tag{13–52}$$

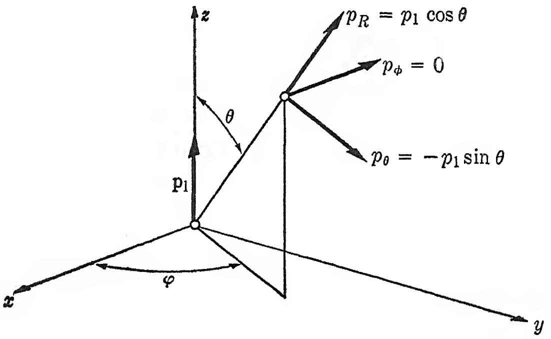


FIG. 13–3 Components of a polarization vector oriented along the polar axis.

It then follows from the expression for the curl in polar coordinates (Appendix II) that C_φ is the only component of $\mathbf{C}$:

$$C_\varphi = \frac{p_1 \sin\theta}{4\pi\epsilon_0 R}\left(\frac{1}{R} - ik\right) e^{ikR}. \tag{13–53}$$

Therefore the magnetic field is just

$$H_\phi = -\frac{i\omega}{4\pi} p_1 \sin\theta \left(\frac{1}{R^2} - \frac{ik}{R}\right) e^{ikR}, \tag{13–54}$$

and the components of the electric field, from Eq. (13–46), are

$$E_R = \frac{1}{R\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta C_\varphi) = \frac{p_1\cos\theta}{2\pi\epsilon_0 R^2}\left(\frac{1}{R} - ik\right) e^{ikR}, \tag{13–55}$$

$$E_\theta = -\frac{1}{R}\frac{\partial}{\partial R}(RC_\varphi) = \frac{p_1\sin\theta}{4\pi\epsilon_0 R}\left(\frac{1}{R^2} - \frac{ik}{R} - k^2\right) e^{ikR}. \tag{13–56}$$

The two terms of Eq. (13–54) represent the induction field and the radiation field. The θ-component of the electric field has three terms: the first is the static dipole field, varying as the inverse cube of the distance from the dipole; the second term varies as the inverse square of the distance and is called the transition field. The transition field will not contribute to the radiated energy but it does contribute to the energy storage during the oscillation. The radiation fields alone are simply

$$H_\phi = -\frac{\omega k p_1 \sin\theta e^{ikR}}{4\pi R}, \tag{13–57}$$

$$E_\theta = -\frac{k^2 p_1 \sin\theta e^{ikR}}{4\pi\epsilon_0 R}. \tag{13–58}$$

In vector form, the radiation fields become

$$\mathbf{E} = \frac{e^{ikR}}{4\pi\epsilon_0 R}[(\mathbf{p}_1 \times \mathbf{k}) \times \mathbf{k}], \quad \mathbf{H} = \frac{\omega e^{ikR}}{4\pi R}(\mathbf{p}_1 \times \mathbf{k}), \tag{13–59}$$

where $\mathbf{k}$ is, of course, along $\mathbf{R}$.

The equations for the electric field lines, or lines of force, can be expressed very simply in terms of $\mathbf{C}$. Since a line of force is defined by the condition that its line element $d\mathbf{s}$ be in the direction of $\mathbf{E}$, the equation of the field lines is just $d\mathbf{s} \times \mathbf{E} = d\mathbf{s} \times (\nabla \times \mathbf{C}) = 0$. In this case, the equation is

$$\frac{\partial}{\partial R}(RC\sin\theta)\, dR + \frac{\partial}{\partial\theta}(RC\sin\theta)\, d\theta = 0, \tag{13–60}$$

AGRICULTURAL

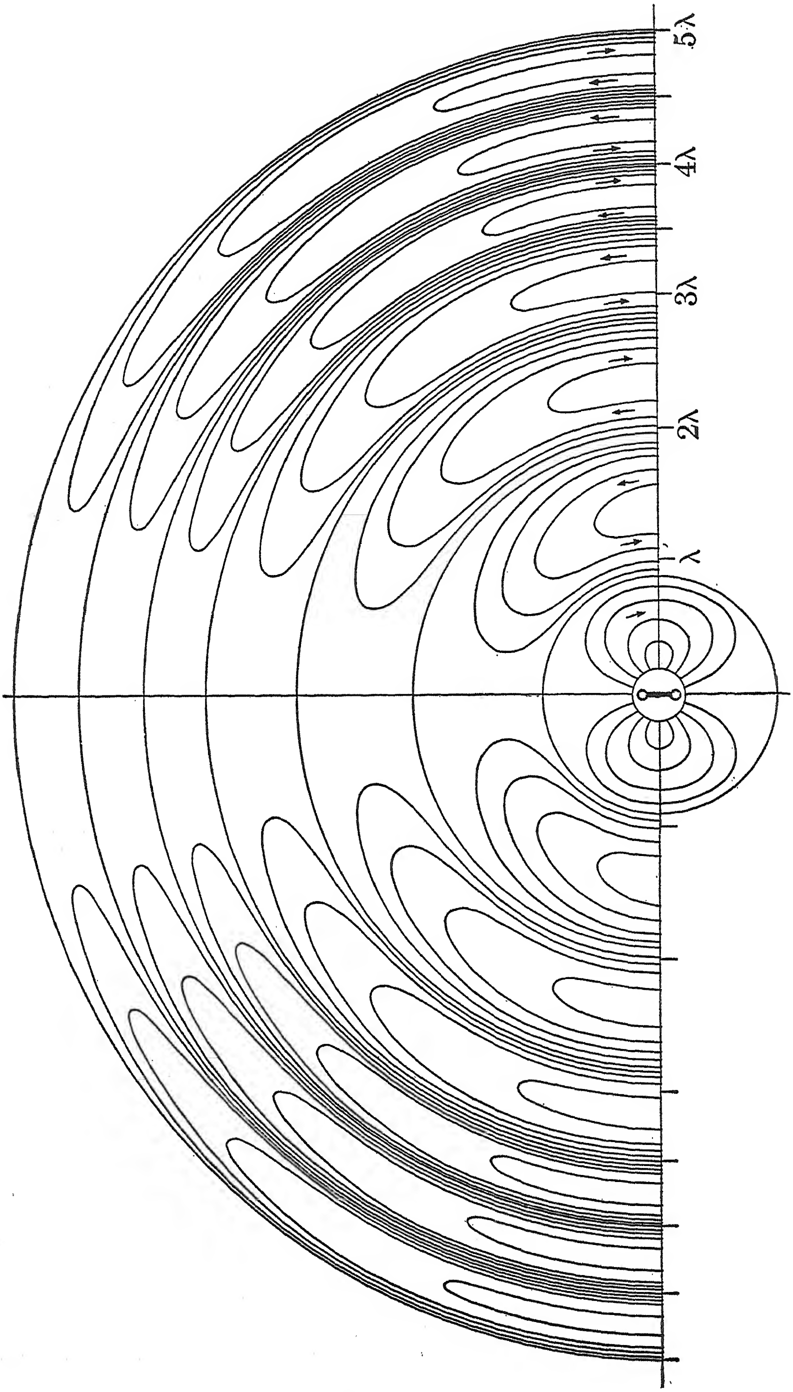


FIG. 13–4 Electric field lines produced by an oscillating dipole.

of which a solution is

$$RC \sin\theta = \text{constant.}$$

On substituting the expression for C from Eq. (13–53), we obtain

$$R^{-1}\sin^2\theta\cos(kR - \omega t - \tan^{-1}kR) = \text{constant} \qquad (13\text{–}61)$$

when the time-dependent term is included. This is an exact equation applying both in the radiation zone and in the induction zone. If we consider the radiation field alone, we find

$$\sin^2\theta\cos(kR - \omega t) = \text{constant}$$

as the equation for the lines of force. This field is periodic radially, with radial spacing equal to the wavelength λ. For smaller values of kR the apparent phase velocity of the wave front, in accord with Eq. (13–61), is greater than c. The transition is indicated in Fig. 13–4, with the regularly spaced lines of force in the radiation field near the periphery. Accurate scale has been sacrificed in order that both fields may be shown in the same drawing; a famous series of diagrams exhibiting the way in which the electric lines of force are initiated in time is to be found in *Gesammelte Werke*, Band II, of H. Hertz, or in the translated volume entitled *Electric Waves.*

13–7 Multipole radiation. Let us now consider the significance of higher moments that contribute to the $\mathbf{\Pi}$ vector. The term $n = 1$ in the general expansion of Eq. (13–50),

$$\Pi_{1\omega} = \frac{e^{ikR}}{4\pi\epsilon_0 R}\int k\xi \mathbf{p}\cos\theta\, dv', \qquad (13\text{–}62)$$

can be written as

$$\Pi_{1\omega} = \frac{ke^{ikR}}{4\pi\epsilon_0 R^2}\int \mathbf{p}(\boldsymbol{\xi}\cdot\mathbf{R})\, dv'. \qquad (13\text{–}63)$$

The components of $\boldsymbol{\xi}$ are x'_α, and those of $\mathbf{R}$ are x_α. Then in tensor notation the integrand of Eq. (13–63) is $x_\alpha p_\beta x'_\alpha$, which is a function of two sets of parameters of the distribution, p_β and x'_α. It is convenient to break up this tensor into two parts, one symmetric and the other antisymmetric in the two parameters. That is, we put

$$x_\alpha p_\beta x'_\alpha = \frac{x_\alpha(p_\beta x'_\alpha + p_\alpha x'_\beta) + x_\alpha(p_\beta x'_\alpha - p_\alpha x'_\beta)}{2}. \qquad (13\text{–}64)$$

This process, much the same as that employed in Section 7–11, is analogous to the process often used in the mechanics of continua, namely, that of separating a general strain of an elastic solid into the sum of a pure strain represented by a symmetric tensor and a body rotation represented by an antisymmetric tensor. Putting the antisymmetric part of Eq. (13–64) back into vector notation, we may write the integral in Eq. (13–63) as

$$\tfrac{1}{2}\int \{[\mathbf{R} \times (\mathbf{p} \times \boldsymbol{\xi})]_\beta + x_\alpha(p_\beta x'_\alpha + p_\alpha x'_\beta)\}\, dv'. \tag{13–65}$$

Let us consider the antisymmetric term first. Since $\mathbf{R}$ is not a function of the primed variables of integration, it can be taken outside the integral sign, and the first term of expression (13–65) is just

$$\tfrac{1}{2}\mathbf{R} \times \int (\mathbf{p} \times \boldsymbol{\xi})\, dv'. \tag{13–66}$$

The significance of this integral can be easily recognized if we express the polarization vector in terms of the current. Since $\mathbf{p} = -\mathbf{j}/i\omega$ for sinusoidal time variation, (13–66) becomes

$$\frac{-\mathbf{R}}{2i\omega} \times \int \mathbf{j} \times \boldsymbol{\xi}\, dv'. \tag{13–67}$$

But the magnetic moment of a current distribution is

$$\mathbf{m} = \tfrac{1}{2}\int \boldsymbol{\xi} \times \mathbf{j}\, dv',$$

and hence (13–67) is given by

$$\frac{1}{i\omega}\mathbf{R} \times \mathbf{m}.$$

The polarization potential $\boldsymbol{\Pi}$ corresponding to the antisymmetric part of its original integrand is therefore

$$\boldsymbol{\Pi}_{\text{antisym.}} = \frac{k e^{ikR}(\mathbf{R} \times \mathbf{m})}{4\pi\epsilon_0 R^2 i\omega}. \tag{13–68}$$

The field can be computed readily from $\boldsymbol{\Pi}$ as given by Eq. (13–68). If we omit the induction terms, we obtain

$$\mathbf{B}_{\text{antisym.}} = \frac{1}{c^2}\frac{\partial \mathbf{C}}{\partial t} = \frac{i\mu_0 k^2[\mathbf{R} \times (\mathbf{m} \times \mathbf{R})]e^{ikR}}{4\pi R^3}, \tag{13–69}$$

which, if we ignore the phase factor, is equivalent to

$$\mathbf{H}_{\text{antisym.}} = \frac{e^{ikR}}{4\pi R}[\mathbf{k} \times (\mathbf{m} \times \mathbf{k})]. \tag{13–70}$$

The magnetic field of Eq. (13–70) has exactly the same mathematical form as that of the electric field in Eq. (13–59) for the electric dipole. It is known as the magnetic dipole radiation field. Magnetic dipole radiation corresponds to a current distribution which, although it has no net oscillating electric dipole, does have a sinusoidally varying circulation of the charges.

Now let us consider the physical significance of the symmetrical term of Eq. (13–64). The corresponding Hertz vector is given by

$$\Pi_{\beta\,\mathrm{sym}} = \frac{ke^{ikR}}{4\pi\epsilon_0 R^2}\, x_\alpha \int \tfrac{1}{2}(p_\beta x'_\alpha + p_\alpha x'_\beta)\, dv'. \tag{13–71}$$

The symmetry of Eq. (13–71) will become more evident if we write it in terms of the charge density $\rho(x'_\alpha)$ and the coordinates of the charge density. Consider a quantity defined by

$$Q_{\alpha\beta} = \int \rho x'_\alpha x'_\beta \, dv' = -\int \frac{\partial p_\gamma}{\partial x'_\gamma}\, x'_\alpha x'_\beta \, dv', \tag{13–72}$$

which is known as the electric quadrupole moment of the charge distribution. If we integrate Eq. (13–72) by parts, and extend the volume of integration outside the charge distribution so that the integrated parts vanish, we find that in terms of the quadrupole moment Eq. (13–71) can be written

$$\Pi_\beta = \frac{ke^{ikR}}{8\pi\epsilon_0 R^2}\, x_\alpha Q_{\alpha\beta}. \tag{13–73}$$

Since the quadrupole moment is represented by the symmetric matrix of Eq. (13–72), it can also be represented by a family of quadrics derived from the quadratic form

$$x_\alpha x_\beta Q_{\alpha\beta} = C_2 = \text{constant}. \tag{13–74}$$

In terms of the parameter C_2 of this equation Π can therefore be written as

$$\Pi = \frac{ke^{ikR}}{8\pi\epsilon_0 R^2}\, \nabla C_2, \tag{13–75}$$

indicating that the direction of Π is everywhere normal to the quadric surfaces of Eq. (13–74).

Let us calculate the components of the fields that correspond to a general quadrupole. We may choose a system of axes x, y, z, along the principal axes of the quadrupole quadric. The components of Π are then given by

$$\begin{Bmatrix} \Pi_x \\ \Pi_y \\ \Pi_z \end{Bmatrix} = \frac{ke^{ikR}}{8\pi\epsilon_0 R} \begin{Bmatrix} \sin\theta\cos\varphi\, Q_{xx} \\ \sin\theta\sin\varphi\, Q_{yy} \\ \cos\theta\, Q_{zz} \end{Bmatrix}. \tag{13–76}$$

The components of **C**, obtained by taking the curl of **Π**, are

$$\begin{aligned} C_\theta &= \frac{ik^2e^{ikR}}{16\pi\epsilon_0 R}\sin\theta\sin 2\varphi(Q_{xx}-Q_{yy}), \\ C_\varphi &= \frac{ik^2e^{ikR}}{32\pi\epsilon_0 R}\sin 2\theta[Q_{xx}+Q_{yy}-2Q_{zz}-(Q_{yy}-Q_{xx})\cos 2\varphi]. \end{aligned} \tag{13–77}$$

And from **C** the electric field components are found to be

$$\begin{aligned} E_\theta &= \frac{k^3e^{ikR}}{32\pi\epsilon_0 R}\sin 2\theta[Q_{xx}+Q_{yy}-2Q_{zz}-(Q_{yy}-Q_{xx})\cos 2\varphi], \\ E_\varphi &= \frac{k^3e^{ikR}}{16\pi\epsilon_0 R}\sin\theta\sin 2\varphi(Q_{yy}-Q_{xx}), \end{aligned} \tag{13–78}$$

while the magnetic field components are

$$\begin{aligned} B_\theta &= -\frac{\omega\mu_0 k^2e^{ikR}}{16\pi R}\sin\theta\sin 2\varphi(Q_{yy}-Q_{xx}), \\ B_\varphi &= \frac{\omega\mu_0 k^2e^{ikR}}{32\pi R}\sin 2\theta[Q_{xx}+Q_{yy}-2Q_{zz}-(Q_{yy}-Q_{xx})\cos 2\varphi]. \end{aligned} \tag{13–79}$$

Note that the fields depend only on the differences in the quadrupole moments; hence five, not six, second moments of the distribution specify the radiation field. Only the radiation fields are given by Eqs. (13–78) and (13–79), and there are no radial components.

Two features of the quadrupole radiation fields can be noted by inspection. First: in case two of the moments are equal, i.e., if the quadrupole is a spheroidal distribution, and if the polar axis of the spheroid is taken as the z-axis, then the only component of **C** and hence of the magnetic field is

$$C_\varphi = \frac{ik^2e^{ikR}}{16\pi\epsilon_0 R}\sin 2\theta(Q_{xx}-Q_{zz}). \tag{13–80}$$

Second: in general the distribution will have two nodal cones where there is zero field, compared with a single nodal line in the case of a dipole distribution. Physically, the quadrupole moment arises from a pulsating charge distribution of such symmetry that the dipole moment remains zero during the pulsation. The simplest example of a quadrupole consists of two dipoles displaced a slight distance from each other and oscillating in opposition.

The formalism of this chapter is of general applicability, but in the detailed application of the retarded potentials to compute the radiation from elementary charges a number of difficulties arise. Some of these problems were clarified in a very simple way by the theory of relativity. We shall see why this theory was necessary before proceeding further with the theory of radiation.

SUGGESTED REFERENCES

J. A. Stratton, *Electromagnetic Theory.* Chapter VIII on radiation includes the retarded potential solution and multipole expansion of the Hertz solution.

P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* includes much highly useful material on the solution of the inhomogeneous wave equation.

G. A. Schott, *Electromagnetic Radiation* contains many examples.

EXERCISES

1. Show that a sphere charged in spherical symmetry and oscillating purely radially will not radiate.

2. Show that if we take the solution of the wave equation as

$$\phi = (\phi_{\text{retarded}} + \phi_{\text{advanced}})/2,$$

$$\mathbf{A} = (\mathbf{A}_{\text{retarded}} + \mathbf{A}_{\text{advanced}})/2,$$

we obtain zero for the total energy radiated by an oscillating system.

3. The current in a circular loop of radius a is given by

$$I = I_0 \sin(n\varphi) e^{-i\omega t}.$$

Show that, except for a phase factor, the vector potential in the radiation zone is

$$A_\varphi = \frac{\mu_0 I_0}{2} \frac{a}{r} J_n'(ka \sin\varphi) \sin n\varphi\, e^{i(kr-\omega t)}.$$

4. An electric dipole lies in the xy-plane at the origin and rotates about the z-axis with constant angular velocity. Find the cylindrical coordinate components of the magnetic field intensity for large distances from the dipole.

5. Consider two dipoles oriented at right angles to each other and oscillating 90° out of phase but at the same frequency. (a) Calculate the energy U emitted in a time dt. (b) Consider a spherical absorber at a large distance R. The angular momentum imparted to this screen in time dt is

$$\mathbf{G} = \left[\int \frac{\mathbf{r} \times (\mathbf{E} \times \mathbf{H})}{c^2}\, dS \right] c\, dt,$$

where the integral is to be taken over the surface of the sphere. Calculate $\mathbf{G}$ and compare with the result of (a). Use the complete dipole fields.

6. A linear quadrupole oscillator consists of charges $-e$, $+2e$, $-e$, with the positive charge stationary at the origin and the negative charges at z_1 and z_2 given by

$$z_1 = -z_2 = a \cos \tfrac{1}{2}\omega t.$$

Compute the fields at large distances and find the average rate at which energy is radiated.

7. Instead of oscillating, the two negative charges of problem 6 rotate with constant angular velocity about the positive charge at the origin, maintaining a fixed distance a from $+2e$ on the line of the three charges. Find the components of the quadrupole moment and the average rate of radiation.

8. Consider a center-fed antenna of total length $L = \lambda/4$ with a current distribution

$$J = J_0 \cos \frac{2\pi\xi}{\lambda}, \qquad -\lambda/8 < \xi < \lambda/8.$$

Calculate the radiated power corresponding to the linear electric dipole and octupole moments and add. Compare with the exact result.

CHAPTER 14

THE EXPERIMENTAL BASIS FOR THE THEORY OF SPECIAL RELATIVITY

14–1 Galilean relativity and electrodynamics. We have found that electromagnetic fields are propagated *in vacuo* with a velocity $c = 1/\sqrt{\mu_0\epsilon_0}$ which is *per se* a characteristic constant of the theory. This is a feature of Maxwell's equations that is at variance with the laws of classical mechanics. Classical mechanics contains no characteristic constants and can be scaled with respect to all physical quantities; Maxwell's equations can be scaled in relation to length and time individually but not as to velocity. This means that Maxwell's theory did not share a noteworthy property of Newtonian mechanics, namely, its independence of any particular frame of reference. Uniform motion is indistinguishable from no motion at all—the laws of motion hold equally in all rigid coordinate frames moving with uniform velocity with respect to each other. Mathematically this can be stated by noting that if a set of primed coordinates is moving with a velocity v along the x-axis of the plane coordinates of Fig. 14–1, substitution of the coordinate relations

$$x' = x - vt, \quad y' = y, \quad z' = z \tag{14–1}$$

leaves the equations of motion unchanged. This can be seen by direct substitution in the basic equation of motion for n mass points whose interaction is describable by means of potential functions depending on their separation. For the ith particle, this equation is

$$m_i\ddot{\mathbf{r}}_i = -\sum_{j\neq i}^{n} \nabla_i[V(|\mathbf{r}_i - \mathbf{r}_j|)], \tag{14–2}$$

and it is obvious that Eq. (14–2) would be exactly the same in the primed coordinates. Equations (14–1) are known as the Galilean transformation,

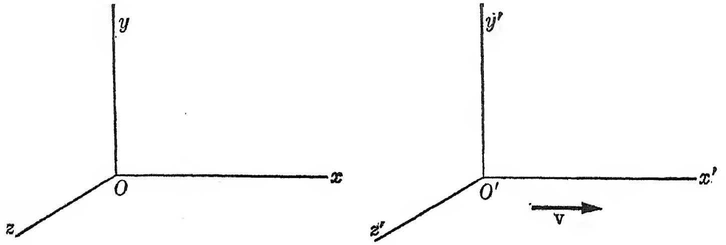


FIG. 14–1 Primed frame moving with respect to unprimed coordinate frame.

and it can thus be said that the laws of classical mechanics are invariant under a Galilean transformation. A coordinate frame in which a body on which no forces are acting is unaccelerated is called an "inertial" frame; we can therefore formulate the invariance of the laws of mechanics under the transformation of Eqs. (14–1) by saying that in classical mechanics all inertial frames are equivalent. This statement is sometimes known as the principle of Galilean relativity.

Since the absence of forces can be detected only by the absence of acceleration unless the sources of force are known, an inertial frame is not strictly definable in a field theory. But there is a more immediate difficulty with Maxwell's equations: the form of the wave equation is not preserved by the substitution of Eqs. (14–1); electromagnetic effects will therefore presumably not be the same if observed from different frames moving with a constant velocity relative to one another. Specifically, the velocity of propagation of a plane wave *in vacuo* would not retain its value $c = 1/\sqrt{\mu_0\epsilon_0}$. If we accept the basic correctness of Maxwell's equations and classical kinematic laws, it follows that there exists a unique privileged frame of reference, the classical "ether frame," in which Maxwell's equations are valid and in which light is propagated with the velocity c.

Since the principle of Galilean relativity does apply to the laws of mechanics but does not apply to electrodynamics, we are forced to choose among the following alternatives:

(a) A principle of relativity exists for mechanics, but not for electrodynamics. A preferred inertial frame (the ether frame) exists in electrodynamics.

(b) A principle of relativity exists for both mechanics and electrodynamics, but electrodynamics is not correct in the Maxwell formulation.

(c) A principle of relativity exists for both mechanics and electrodynamics, but the laws of mechanics in the Newtonian form need modification.

The choice between these possibilities can be made only on the basis of experimental results. We shall see by analysis of the relevant experiments that alternative (c), in the form of the special theory of relativity, is essentially correct. A reference list of the basic experiments is given at the end of the chapter, and further references may be found in the more comprehensive treatments of relativity also listed. The fundamental experiments fall into three main classes: (a) attempts to locate a preferred inertial frame for the laws of electrodynamics; (b) attempts to obtain deviations from the laws of classical electrodynamics; and (c) attempts to observe deviations from classical mechanics.

14–2 The search for an absolute ether frame. It can be seen in an elementary way that two charges, q and $-q$, fixed at the ends of a rigid rod and set in motion with velocity $\mathbf{v}$, would affect each other in the same way

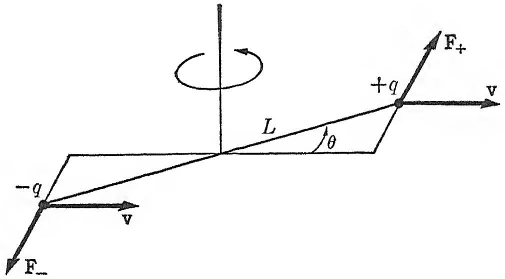


FIG. 14–2 Equal and opposite charges moving with velocity $\mathbf{v}$. $\mathbf{F}_+$ and $\mathbf{F}_-$ are not collinear.

as two current elements of magnitude $J\,d\mathbf{l} = q\mathbf{v}$. The forces on the two current elements would be directed oppositely, but in general they would not be collinear. Under these conditions the rod, if free to turn, would tend to set itself at right angles to the velocity. It is of interest to compute the size of this effect. By Ampere's law (with reference to Fig. 14–2),

$$\mathbf{F} = \frac{\mu_0}{4\pi} q^2 \frac{\mathbf{v} \times (\mathbf{v} \times \mathbf{L})}{L^3}, \tag{14–3}$$

where $\mathbf{L}$ represents both the length and the direction of the rod. The magnitude of the force given by Eq. (14–3) may be written

$$F = \frac{1}{4\pi\epsilon_0} \frac{q^2}{L^2} \frac{v^2}{c^2} \sin\theta, \tag{14–4}$$

and its direction is perpendicular to $\mathbf{v}$ in the plane of $\mathbf{L}$ and $\mathbf{v}$. Equation (14–4) indicates that the effect is of order $(v/c)^2$ in comparison with the electrostatic interaction of the two charges. The corresponding value of the torque, i.e., the net value of the couple shown in Fig. 14–2, is

$$\begin{aligned} FL\cos\theta &= \frac{1}{4\pi\epsilon_0} \frac{q^2v^2 \sin\theta\cos\theta}{c^2L} \\ &= \frac{1}{8\pi\epsilon_0} \frac{q^2}{L} \frac{v^2}{c^2} \sin 2\theta. \end{aligned} \tag{14–5}$$

Now if there is a preferred frame of coordinates it seems very unlikely that it is the frame with respect to which the earth is at rest, and the translatory motion of the earth should produce such a torque. An experiment involving a delicately suspended parallel plate condenser was performed by Trouton and Noble, and repeated later with even greater accuracy, but it failed to give any indication of a torque, even though its magnitude with v equal to the known velocity of the earth in its orbit should be easily measurable.

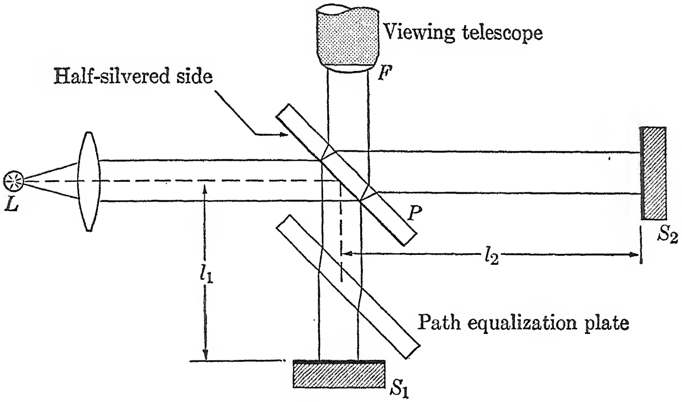


FIG. 14-3 The Michelson-Morley experiment.

The most famous attempt to localize the ether frame was the Michelson-Morley experiment. Light from a source L in Fig. 14-3 is split into two beams by a half-silvered mirror at P. The beams are reflected at mirrors S_1 and S_2 respectively and return through the half-silvered mirror to the telescope at F, where interference fringes are observed. Let us assume that relative to a stationary ether the instrument is moving with a velocity v parallel to S_1P. By the arguments of classical physics, the time required for light to traverse the path PS_1P is

$$t_1 = l_1\left(\frac{1}{c-v} + \frac{1}{c+v}\right) = \frac{2l_1}{c(1-\beta^2)}, \tag{14-6}$$

where $\beta = v/c$. In computing the time required for the path PS_2P, we must take account of the fact that P will move through the distance δ (see Fig. 14-4) while the light travels from P to S_2. δ is given by

$$\frac{\delta}{\sqrt{\delta^2 + l_2^2}} = \frac{v}{c}; \quad \delta = \frac{\beta l_2}{\sqrt{1-\beta^2}}. \tag{14-7}$$

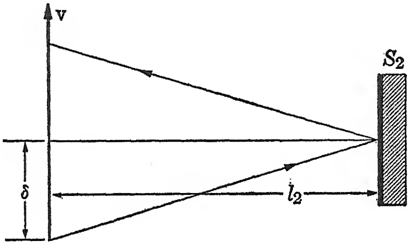


FIG. 14-4 Detail of the Michelson-Morley experiment in the calculation of **δ**.

Hence, for path PS_2P,

$$t_2 = \frac{2}{c}\sqrt{l_2^2 + \delta^2} = \frac{2l_2}{c\sqrt{1-\beta^2}}. \tag{14-8}$$

The difference Δ in optical path is

$$c(t_1 - t_2) = \frac{2}{\sqrt{1-\beta^2}}\left(\frac{l_1}{\sqrt{1-\beta^2}} - l_2\right). \tag{14-9}$$

If the instrument is rotated through 90°, l_1 and l_2 interchange roles, and

$$t_1' = \frac{2l_1/c}{\sqrt{1-\beta^2}}, \quad t_2' = \frac{2l_2/c}{1-\beta^2}.$$

The new difference in optical path is

$$\Delta' = c(t_1' - t_2') = \frac{2}{\sqrt{1-\beta^2}}\left(l_1 - \frac{l_2}{\sqrt{1-\beta^2}}\right).$$

Thus, on rotating the apparatus through 90°, we should expect the interference pattern to shift by n fringes, where n is given by

$$n = \frac{\Delta' - \Delta}{\lambda} = \frac{2(l_1 + l_2)}{\lambda\sqrt{1-\beta^2}}\left(1 - \frac{1}{\sqrt{1-\beta^2}}\right)$$

$$\simeq -\frac{(l_1 + l_2)}{\lambda}\beta^2, \tag{14-10}$$

since v is small compared with c.

No such shift was observed by Michelson and Morley. The estimated accuracy of their result was 10 km/sec, i.e., the velocity of the earth relative to any "ether frame" must be less than 10 km/sec, although the velocity of the earth in its orbit is roughly 30 km/sec. Much discussion resulted from the experiments of Miller, which gave positive indication of a velocity of about 10 km/sec apparently directed toward a certain point in space, but improved technique has confirmed the absence of a fringe shift, and a new analysis of Miller's data has shown that they are not inconsistent with those of other observers. A summary of the various trials of the Michelson-Morley experiment is given in Table 14–1, to which is appended a list of the appropriate journal references.* In these experi-

---

* "New Analysis of the Interferometer Observations of Dayton C. Miller," by R. S. Shankland, S. W. McCuskey, F. C. Leone, and G. Kuerti, *Rev. Modern Phys.* **27**, 167 (1955). Table 14–1 is reproduced here through the courtesy of these authors.

Table 14–1

TRIALS OF THE MICHELSON-MORLEY EXPERIMENT

Observer	Year	Place	l	$(2l/\lambda)(v/c)^2$	A	Ratio
Michelson(a)	1881	Potsdam	120 cm	0.04 fringe	0.01 fringe	2
Michelson & Morley(b)	1887	Cleveland	1100	0.40	0.005	40
Morley & Miller(c)	1902–04	Cleveland	3220	1.13	0.0073	80
Miller(d)	1921	Mt. Wilson	3200	1.12	0.04	15
Miller(e)	1923–24	Cleveland	3200	1.12	0.015	40
Miller(f) (sunlight)	1924	Cleveland	3200	1.12	0.007	80
Tomaschek(g) (starlight)	1924	Heidelberg	860	0.3	0.01	15
Miller(h)	1925–26	Mt. Wilson	3200	1.12	0.044	13
Kennedy(i)	1926	Pasadena & Mt. Wilson	200	0.07	0.001	35
Illingworth(j)	1927	Pasadena	200	0.07	0.0002	175
Piccard & Stahel(k)	1927	Mt. Rigi	280	0.13	0.003	20
Michelson *et al.*(l)	1929	Mt. Wilson	2590	0.9	0.005	90
Joos(m)	1930	Jena	2100	0.75	0.001	375

(a) A. A. Michelson, *Am. J. Sci.* **22,** 120 (1881); *Phil. Mag.* **13,** 236 (1882).

(b) A. A. Michelson and E. W. Morley, *Am. J. Sci.* **34,** 333 (1887); *Phil. Mag.* **24,** 449 (1887).

(c) E. W. Morley and D. C. Miller, *Phil. Mag.* **9,** 680 (1905); *Proc. Am. Acad. Arts Sci.* **41,** 321 (1905).

(d) D. C. Miller, *Data sheets of observations,* Dec. 9–11, 1921 (unpublished).

(e) D. C. Miller, *Observations,* Aug. 23–Sept. 4, 1923; June 27–July 26, 1924 (unpublished).

(f) D. C. Miller, Observations with Sunlight on July 8–9, 1924, *Proc. Nat. Acad. Sci.* **11,** 311 (1925).

(g) R. Tomaschek, *Ann. Physik* **73,** 105 (1924).

(h) D. C. Miller, *Rev. Mod. Phys.* **5,** 203 (1933).

(i) R. J. Kennedy, *Proc. Nat. Acad. Sci.* **12,** 621 (1926); *Astrophys. J.* **68,** 367 (1928).

(j) K. K. Illingworth, *Phys. Rev.* **30,** 692 (1927).

(k) A. Piccard and E. Stahel, *Compt. rend.* **183,** 420 (1926); **184,** 152, 451 (1927).

(l) A. A. Michelson, F. G. Pease, and F. Pearson, *Nature* **123,** 88 (1929); *J. Opt. Soc. Am.* **18,** 181 (1929).

(m) G. Joos, *Ann. Physik* **7,** 385 (1930); *Naturwiss.* **38,** 784 (1931).

ments the interferometer arms were equal, and $(2l/\lambda)(v/c)^2$, by the ether theory, is the shift expected due to the earth's motion. The amplitude of the second harmonic of the fringe shift actually found by each observer is given in the column headed A. To provide a comparison of accuracy of the various trials, the last column gives the ratio of the shift expected prerelativistically to $2A$. It is clear that a null result can be accepted with confidence.

14–3 The Lorentz-Fitzgerald contraction hypothesis. We have seen that the search for a preferred frame for electrodynamics without further modification of either electrodynamics or mechanics was unsuccessful. An attempt to preserve the concept of the preferred ether frame despite the negative result of the Michelson-Morley experiment led to the contraction hypothesis. Lorentz and Fitzgerald postulated that as a result of motion relative to the stationary ether all bodies are contracted by the factor $\sqrt{1-\beta^2}$ in the direction of motion. On this hypothesis, l_1 in Fig. 14–3 is actually equal to $l_1^0\sqrt{1-\beta^2}$, where l_1^0 is the length of l_1 when at rest with respect to the ether, while $l_2 = l_2^0$. Hence, Eq. (14–9) becomes

$$\Delta = \frac{2}{\sqrt{1-\beta^2}}(l_1^0 - l_2^0) \tag{14–11}$$

and no fringe shift would be obtained by rotation through 90°. Furthermore, if $l_1^0 \simeq l_2^0$, as was the case in the original experiments, no fringe shift would occur as a result of changing the velocity. If $l_1^0 \neq l_2^0$, however, even with the Lorentz contraction a fringe shift given by

$$n = \frac{l_1^0 - l_2^0}{\lambda}(\beta'^2 - \beta^2) \tag{14–12}$$

would be expected from a velocity change, as indicated by the factor $(\beta'^2 - \beta^2)$.

An interferometer using a path difference essentially as long as coherence of the source permitted was constructed by Kennedy. The square of the velocity of the instrument is presumably given by

$$v^2 = c^2\beta^2 = (\mathbf{v}_E + \mathbf{v}_R + \mathbf{v}_S)^2,$$

where $\mathbf{v}_E$ is the velocity of the earth with respect to the sun, $\mathbf{v}_R$ is the surface velocity of the earth due to its rotation, and $\mathbf{v}_S$ is the velocity of the sun. Every twelve hours this quantity should change by

$$\Delta v^2 = 4(\mathbf{v}_S + \mathbf{v}_E)\cdot\mathbf{v}_R,$$

and every six months it should change by

$$\Delta v^2 = 4(\mathbf{v}_S + \mathbf{v}_R)\cdot\mathbf{v}_E.$$

Neither effect was observed, in contradiction to the Lorentz contraction hypothesis.

14–4 "Ether drag." A further alternative in which the concept of the ether could be reconciled with the Michelson-Morley result was to consider the ether frame attached to ponderable bodies. This would automatically give a null result for terrestrial interferometer experiments. The

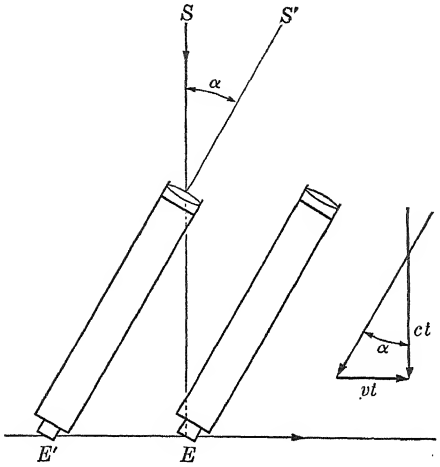


Fig. 14-5 Aberration of starlight. *SE* is the actual path of the light from the star at zenith and *S'E'* is the apparent path.

assumption of a local ether, however, is in direct contradiction to two well-established phenomena.

The first is the aberration of "fixed" stars. Due to the motion of the earth about the sun, distant stars appear to move in orbits approximately 41″ in angular diameter. Consider a star at the zenith of the ecliptic. If this star is to be observed through a telescope the telescope tube must be tilted toward the direction of the earth's motion by an angle α, as shown in Fig. 14-5. It is seen from the figure that, classically,

$$\tan \alpha = v/c, \tag{14-13}$$

and with 30 km/sec for the velocity of the earth in its orbit, $\alpha = 10^{-4} = 20.5''$, in agreement with observation. If the ether were dragged by the earth in its motion we should expect no aberration to occur.

The second contradiction arises in connection with the propagation of electromagnetic waves in a moving medium. We have seen in Section 11-4 that only the part of the velocity of light which depends on the polarization current $\partial \mathbf{P}/\partial t$ is affected by the motion of the medium, while the $\epsilon_0(\partial \mathbf{E}/\partial t)$ term remains unchanged. If the ether moved with the medium, the propagation velocity would be simply $c \pm v$, instead of being given by the Fresnel-Fizeau coefficient. This conclusion is contrary to observation. The "ether drag" hypothesis thus leads to discrepancies of the first order in v/c.

These considerations make the idea of a preferred frame appear unacceptable, even when it is only locally stationary. We are therefore led to the alternative that a principle of relativity must be valid in electrodynamics. This is equivalent to demand for modification of either electrodynamics or classical mechanics. We shall first consider the attempts to modify electrodynamics in such a way as to escape the paradoxes associated with the idea of an ether frame.

14–5 Emission theories. In the so-called emission theories electrodynamics is modified by supposing that the velocity of a light wave remains associated with the source rather than with a local or universal frame. In all emission theories it is postulated that the velocity of light is c/n relative to the original source, and independent of the state of the transmitting medium of refractive index n. The theories differ among themselves in predicting what happens to the velocity of light upon reflection from a moving mirror. After a reflection three alternatives are possible:

(1) The velocity remains c/n relative to the original source.

(2) The velocity becomes c/n relative to the mirror (the so-called ballistic theory).

(3) The velocity becomes c/n relative to the mirror image.

The first alternative was proposed by Ritz, and is the only one of the three theories that does not lead to coherence difficulties with reflected light. Ritz retained two of Maxwell's equations:

$$\nabla \cdot \mathbf{B} = 0 \quad \text{and} \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t},$$

with **E** and **B** derived from

$$\mathbf{B} = \nabla \times \mathbf{A}, \quad \mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t},$$

as before. But he replaced the two equations involving sources by

$$\phi = \frac{1}{4\pi\epsilon_0} \int \frac{\rho\left(x'_\alpha, t - \dfrac{r}{c + v_r}\right)}{r}\, dv', \tag{14–14}$$

$$\mathbf{A} = \frac{\mu_0}{4\pi} \int \frac{\mathbf{j}\left(x'_\alpha, t - \dfrac{r}{c + v_r}\right)}{r}\, dv', \tag{14–15}$$

where v_r is the component of the velocity of the source in the direction of the vector joining the source and the point of observation. Equations (14–14) and (14–15) replace the ordinary retarded potential solutions given by Eqs. (13–24) and (13–25). In this way, fields due to a moving source are definable.

In comparison with ordinary electromagnetic theory all three emission theories predict differences of the first order in v/c for experiments on the interference between light beams reflected from moving mirrors. Thomson, Majorana, and Stewart performed such experiments and obtained results in disagreement with the second and third of the emission theories but agreeing, within experimental error, with the predictions of Ritz's theory. This apparent agreement is due to the fact that as applied to any closed system of interfering beams the Ritz theory yields results differing only by terms of the second order in v/c from those to be expected if the velocity of light is constant. According to Ritz, if the source is moving toward the mirror with velocity v the time required to traverse a given distance l on the forward trip is $l/(c + v)$, while on the return trip, after reflection from the mirror, it is $l/(c - v)$. For the total elapsed time this gives

$$\Delta t = \frac{2l}{c(1 - \beta^2)},$$

which differs only by terms in the second order from the expression for constant velocity of light, $t = 2l/c$. Thus terrestrial moving source and mirror experiments fail to give a first order contradiction to the Ritz emission theory.

There are, however, two extraterrestrial phenomena which contradict any form of emission theory. The dynamics of eclipsing binary stars has been carefully analyzed by de Sitter. If the velocity of light depends additively on the velocity of the source, it is evident that the time for light to reach the earth from the approaching star should be smaller than that from the receding member of the doublet. De Sitter showed that this would have the effect of introducing a spurious eccentricity into the orbit as calculated by the laws of mechanics. Actually, no such effect is observed; in fact, de Sitter concluded that if $v_{\text{light}} = c + kv_{\text{star}}$, then $k < 0.002$. The second piece of extraterrestrial evidence is the experience of Miller that the Michelson-Morley experiment is not affected when light from the sun is used. The Ritz theory would predict complications of the interference pattern due, for example, to the sun's rotation.

14-6 Summary. In this chapter we have examined various kinds of experimental evidence for the incompatibility of electrodynamics and Newtonian mechanics. A summary of the relevant experiments, including some whose bearing on the subject will not become evident until we have discussed the theory of relativity, is given in Table 14-2. Their relation to attempts to reconcile them with theory is indicated, and Table 14-3 lists the basic assumptions of the three general alternatives given in Section 14-1. It is clear from Table 14-2 that the experimental basis for the theory of relativity, which modifies Newtonian mechanics, consists essentially of observations in contradiction to all reasonable alternatives.

TABLE 14–2

Theory		Light propagation experiments							Experiments from other fields					
		Aberration	Fizeau convection coefficient	Michelson-Morley	Kennedy-Thorndike	Moving sources and mirrors	De Sitter spectroscopic binaries	Michelson-Morley, using sunlight	Variation of mass with velocity	General mass-energy equivalence	Radiation from moving charges	Meson decay at high velocity	Trouton-Noble	Unipolar induction, using permanent magnet
Ether theories	Stationary ether, no contraction	A	A	D	D	A	A	D	D	N	A	N	D	D
	Stationary ether, Lorentz contraction	A	A	A	D	A	A	A	A	N	A	N	A	D
	Ether attached to ponderable bodies	D	D	A	A	A	A	A	D	N	N	N	A	N
Emission theories	Original source	A	A	A	A	A	D	D	N	N	D	N	N	N
	Ballistic	A	N	A	A	D	D	D	N	N	D	N	N	N
	New source	A	N	A	A	D	D	A	N	N	D	N	N	N
Special theory of relativity		A	A	A	A	A	A	A	A	A	A	A	A	A

Legend: A, the theory agrees with experimental results.
D, the theory disagrees with experimental results.
N, the theory is not applicable to the experiment.

TABLE 14–3

	Emission theory	Classical ether theory	Special theory of relativity
Reference system	No special reference system	Stationary ether is special reference system	No special reference system
Velocity dependence	The velocity of light depends on the motion of the source	The velocity of light is independent of the motion of the source	The velocity of light is independent of the motion
Space-time connection	Space and time are independent	Space and time are independent	Space and time are interdependent
Transformation equations	Inertial frames in relative motion are connected by a Galilean transformation	Inertial frames in relative motion are connected by a Galilean transformation	Inertial frames in relative motion are connected by a Lorentz transformation

There is no single experiment that proves relativity theory. The experiments do present evidence that:

(1) The existence of an ether, either stationary or carried convectively, is undemonstrable.

(2) Modifications of electrodynamics, like the emission theories, are untenable.

It is then plausible to conclude that the basic laws of mechanics need modification.

In 1905, compatible with the experimental facts known at that time, Einstein proposed the following postulates as a solution:

1. The laws of electrodynamics (including, of course, the propagation of light with the velocity c in free space), as well as the laws of mechanics, are the same in all inertial frames.

2. It is impossible to devise an experiment defining a state of absolute motion, or to determine for any physical phenomena a preferred inertial frame having special properties.

If the laws of physics conform to these two postulates all the experiments listed in Table 14–2 are explicable. In the next chapter we shall begin to examine the implications of these postulates, known as the postulates of special relativity.

ADDITIONAL PAPERS ON THE EXPERIMENTAL BASIS OF SPECIAL RELATIVITY

1. TROUTON AND NOBLE, *Phil. Trans.* **A202,** 165 (1903); *Proc. Roy. Soc.* (London) **72,** 132 (1903). Reporting attempts to find torque on a charged condenser.
2. LORENTZ, "Versuch einer Theorie der elektrischen und optischen Erscheinungen in bewegten Korpern," Leiden, 1895. The sections on the contraction hypothesis are reprinted in *The Principle of Relativity,* listed below.
3. KENNEDY AND THORNDIKE, *Phys. Rev.* **42,** 400 (1932). Interferometer with unequal arms. Null result to ±10 km/sec.
4. RITZ, *Ann. Chim. et Phys.* **13,** 145 (1908). Original source emission theory.
5. TOLMAN, *Phys. Rev.* **31,** 26 (1910); *Phys. Rev.* **35,** 136 (1912). THOMSON, J. J., *Phil. Mag.* **19,** 301 (1910). STEWART, *Phys. Rev.* **32,** 418 (1911). Discussions of various emission theories.
6. COMSTOCK, *Phys. Rev.* **30,** 267 (1910); DE SITTER, *Proc. Amsterdam Acad.* **15,** 1297 (1913), and **16,** 395 (1913). Spectroscopic work on binary stars.
7. MAJORANA, *Phil. Mag.* **35,** 163 (1918), and **37,** 145 (1919). Moving source and mirror experiments.
8. KENNEDY, *Phys. Rev.* **47,** 965 (1935). Critical discussion of geometrical effects of high order in the Michelson-Morley experiment.
9. IVES, *J. Opt. Soc. Am.* **215,** (1938). Doppler shift.

Further references may be found in the books on relativity theory listed below.

SUGGESTED REFERENCES

R. BECKER, *Theorie der Elektrizität,* Band II. The experimental basis of the special theory of relativity is discussed with great clarity, pp. 255–269 particularly.

P. G. BERGMANN, *An Introduction to the Theory of Relativity.* The first three chapters of this excellent book form a very readable account of the necessity for special relativity.

A. EINSTEIN, H. A. LORENTZ, H. MINKOWSKI, AND H. WEYL, *The Principle of Relativity,* "A Collection of Original Memoirs of the Special and General Theory of Relativity, together with Notes by A. Sommerfeld." Especially appropriate here are the section on "Michelson's Interference Experiment" by Lorentz and parts of his paper on "Electromagnetic Phenomena in a System Moving with any Velocity Less than that of Light."

R. TOLMAN, *Relativity, Thermodynamics, and Cosmology.* Although the experiments are not discussed in detail, a summary is included in Chapter II, together with references to original papers.

C. MØLLER, *The Theory of Relativity.*

EXERCISES

1. Find what the wave equation becomes under a Galilean transformation.

2. The torque given by Eq. (14–5) may be written as $\frac{1}{2}U_{\text{el}}(v/c)^2 \sin 2\theta$, where U_{el} is the electrostatic energy of the system of opposite charges. Show that for a parallel plate condenser, actually used by Trouton and Noble, the torque is given by $U_{\text{el}}(v/c)^2 \sin 2\theta$, differing from Eq. (14–5) by a factor of 2.

3. Find the torque to be expected on a parallel plate condenser suspended at an angle of 45° to the direction of the earth's orbital velocity, 30 km/sec. Take a plate area of 100 cm^2, separation 1 cm, and a charging voltage of 10 kilovolts.

CHAPTER 15

RELATIVISTIC KINEMATICS AND THE LORENTZ TRANSFORMATION

In addition to Einstein's formal postulates, there are two auxiliary principles implicit in the theory of relativity, the recognition of which is important for the necessary revision of physical concepts. The first is "invariance of the sense of time." The asymmetry of experience in time has been the subject of much discussion. From the standpoint of formal physics there is only one concept which is asymmetric in the time, namely, entropy. But this makes it reasonable to assume that the second law of thermodynamics can be used to ascertain the sense of time independently in any frame of reference; i.e., we shall take the positive direction of time to be that of statistically increasing disorder, or increasing entropy and degradation of heat. The second assumption might be called "the invariance of proper quantities," or the "law of reproducibility of proper quantities." By this we mean that, whenever a measuring experiment is performed, the length of a standard (such as the wavelength of a given spectral line or a crystal lattice spacing) and the rate of a fundamental clock (such as a radioactive decay period) shall be the same as seen by an observer at rest relative to the standard. These additional postulates make it possible for us to retain exact and reproducible meaning for length and time in the course of analyzing the consequences of the postulates of special relativity.

15–1 The velocity of light and simultaneity. We have seen in Chapter 14 that there is much experimental evidence for the conclusion that a principle of relativity exists for all fields of physics, including electrodynamics. This implies, among other things, that the velocity of propagation, c, of plane electromagnetic waves in free space must be independent of the observer's inertial frame. Plausible as it may seem, this statement runs grossly contrary to intuition. Consider, for example, a light pulse starting from a point O, and consider the event as recorded by observers stationed in two frames, one frame containing O at the origin, while the other moves relative to O with a velocity v. Let the origins of the two frames coincide at the emission of the pulse. According to the principle of relativity, *both* observers must see the light wave propagating as a spherical wave centered at their *respective* origins! If we hold that the position of the wave front is an event permitting description independently in space and in time, then we cannot accept this statement as true. The independence of the velocity of light of the particular frame therefore requires a revision of the customary practice of specifying the position coordinates of an event with reference to the particular frame while specifying the time of the event on a "universal" time scale. The paradox of the two wave fronts hinges on the assumption that simultaneity is independent of the frame of the observer. If a disagreement as to the simultaneity of passage through a set of points were permitted to exist, then presumably a kinematics could be constructed in which a spherical light wave would be seen in both frames of reference.

Consequently, we are led to re-examine the concept of simultaneity. If we must abandon the existence of a universal time as not corresponding to reality, then we must establish a mechanism whereby simultaneity can be established in a given frame. The mechanism must be such that a measurement of the velocity of light in the particular frame using its time and distance scale must give c. This means that the only way in which simultaneity can be defined is by means of the velocity of light itself. This conclusion gives c a much more fundamental significance than just the velocity of propagation of electromagnetic waves: it introduces c into all the relations of physics. For example, the utilization of c as the defining element of simultaneity precludes the existence of the "ideal rigid body" of mechanics; if there were such a body its ends would move simultaneously as observed from any frame, and it could therefore be used to establish a "universal time," in violation of our former conclusion.

We therefore consider two instants of time t_1 and t_2 observed at two points x_1 and x_2 in a particular frame as simultaneous if a light wave emitted at the geometrically measured mid-point between x_1 and x_2 arrives at x_1 at the time t_1 and at x_2 at the time t_2.

This definition will automatically ensure that a light pulse emitted at the origin will reach all equidistant points simultaneously and that the wave surface is therefore a sphere in a particular reference frame. Simultaneity of two events at two spatially separated points thus has no significance independent of the frame. The relation of the time intervals observed by two different frames will depend on the spatial interval between the events, and the Galilean transformation, Eqs. (14–1), which transformed temporal intervals as observed by two frames independently of spatial coordinates cannot be in agreement with the simultaneity definition in terms of c. We must therefore attempt to derive the corresponding transformations from an $(x,y,z;t)$ frame to an $(x',y',z';t')$ frame which will supersede the Galilean transformation. Such a transformation must remain linear, to assure mathematical equivalence of all points in space and time, but the spatial and temporal coordinates need not transform independently.

15–2 Kinematic relations in special relativity. The transformation we seek will give the relations between the space-time coordinates of an arbitrary event $(x,y,z;t)$ as observed in a coordinate frame we may designate as Σ, and the space-time coordinates $(x',y',z';t')$ in a frame Σ' moving with uniform velocity relative to Σ. The transformation must obey the postulates of special relativity, given in Section 14–6, for an event of any type. To see how the basic postulates of relativity necessitate the nature of the transformation, we can construct a set of "thought experiments" (*Gedanken Experimente*) so as to introduce one feature of the transformation at a time.

EXPERIMENT I. *Comparison of parallel measuring sticks oriented perpendicular to their direction of relative motion* (see Fig. 15–1). It is explicitly assumed that the properties of a given body of specified structure are independent of its past history when observed in a frame where the body is at rest. (This is called the "proper frame" of the body, and the length of a rod when measured in a frame in which the rod is at rest is called its "proper length.") It is therefore possible to demand that the two measuring sticks, OP and $O'P'$, of Fig. 15–1 be of equal proper length, whether they can be brought to rest with respect to each other or not; for instance, it could be specified that the length of each rod should be a given number of wavelengths of a particular spectral line measured in each frame. Both rods are at right angles to $\mathbf{v}$ (assumed along the x-axis), the velocity of the system Σ' as measured in frame Σ. (To an observer in Σ', Σ will have a velocity $-\mathbf{v}$.)

Let the two systems approach each other so that the mid-points of the rods, M and M', will coincide at passing. Considering Σ' to be moving and Σ to be stationary, let light signals be sent from O' and P' when O' and P' coincide with the y-axis. Since $O'M$ remains equal to $P'M$ during the motion, O' and P' will appear to cross the y-axis simultaneously in both systems. Similarly, O and P will cross the y'-axis simultaneously in both systems. Since the time of observation for both ends is defined identically in both frames, both observers can compare the positions of the end markers at the time of crossover and arrive at the same result. Hence *both* observers would conclude either that $OP \leq O'P'$ or $OP \geq O'P'$. Both systems are equivalent in every way, so that an asymmetric relation would provide a means of determining absolute motion, a possibility ruled

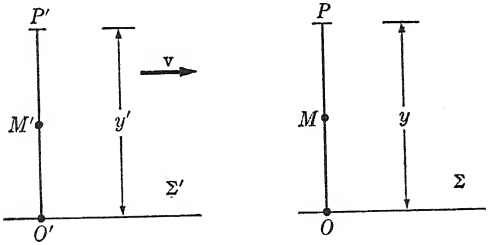


FIG. 15–1 Comparison of parallel measuring sticks oriented perpendicular to their direction of relative motion. $\mathbf{v}$ is the velocity of the Σ' frame measured in Σ.

out by the relativity postulate. We therefore conclude that lengths at right angles to the motion are the same in both frames, and the transformation for such lengths is

$$y' = y, \quad z' = z, \tag{15-1}$$

just as in the Galilean transformation.

EXPERIMENT II. *Comparison of clock rates.* In comparing the rates of clocks attached to coordinate frames in relative motion, we must recognize that it is impossible to compare *one* clock in Σ with *one* clock in Σ', since the clocks will not stay in coincidence. We must compare *two* clocks in Σ with *one* clock in Σ', and have the two clocks in Σ synchronized by means of light signals. Let a source of light be at the same position as the clock in Σ', as shown in Fig. 15-2. A light signal emitted normal to $\mathbf{v}$ is reflected by a mirror which is normal to the z'-axis and at a distance z_0' from the source, and returns to the clock. For an observer at rest in Σ' the time interval between sending and receiving the pulse is

$$\Delta t' = \frac{2z_0'}{c}. \tag{15-2}$$

Observers at rest in Σ can record the time interval Δt between the same two events with two clocks spaced a distance $v\,\Delta t$ apart along x, the direction of the motion. Since c is independent of the frame, this interval will be given by

$$c\,\Delta t = 2\sqrt{(z_0)^2 + \left(\frac{v\,\Delta t}{2}\right)^2},$$

$$\Delta t = \frac{2z_0}{c}\frac{1}{\sqrt{1-\beta^2}}. \tag{15-3}$$

Since $z_0 = z_0'$, it follows that

$$\Delta t = \frac{\Delta t'}{\sqrt{1-\beta^2}}, \tag{15-4}$$

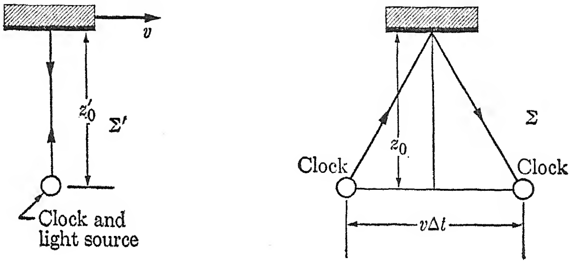


FIG. 15-2 Comparison of clock rates.

where $\Delta t'$ is the proper time interval, i.e., the time interval between two events occurring at the same place in the Σ' frame. Δt is not a proper time interval, for it is measured by two different clocks at different places in Σ. As in the case of a proper length, discussed above, a proper time interval is a definite function of the physical nature of the clock; e.g., a particular radioactive decay constant, or the natural frequency of a crystal of specified proper dimensions, is a constant in the frame where such time intervals are observable at a single point, i.e., in a frame where such a "clock" is at rest.

Note that the observer in Σ will find that his time interval Δt is longer than the proper time interval. This phenomenon is known as *time dilation.* For example, the lifetime of a high velocity meson disintegrating in flight appears lengthened, to a ground observer, by an amount depending on the meson velocity. On the other hand, the lifetime in the proper frame, i.e., the frame at rest with respect to the meson, is an invariant.

EXPERIMENT III. *Comparison of lengths parallel to the direction of motion.* Let us consider a rod of length x_0' in the frame in which it is at rest, i.e., a rod of proper length x_0'. In the Σ frame its length x_0 would be the distance between the ends of the rod observed "simultaneously" in the sense of the simultaneity definition in terms of c. To disentangle the length comparison from the simultaneity calculation, let us consider the following event: a light source S' at one end of the rod in frame Σ' sends a light pulse to a mirror M' at the other end, where it is reflected back to the source. The time interval between the emission and return of the signal is $\Delta t'$. Note that $\Delta t'$ is a proper time interval, being observable with a single clock at one point. Evidently,

$$\Delta t' = \frac{2x_0'}{c}. \tag{15-5}$$

Seen from Σ, these same events appear more complicated. With reference to Fig. 15-3, the source S' was at S_0 at the time of emission, and the mirror M' was at M_0. At the time of reflection, the mirror M' has moved to M, and the pulse returns to the source S' when it is at S_2. The time interval Δt is therefore measured between the points S_0 and S_2 with two clocks, as in Experiment II. Δt is not a proper time, and Eq. (15-4) will apply.

Now, by definition, we mean by x_0 the distance $S_0M_0 = S_1M$. Since M_0 has moved to M with velocity v while the light moved from S_0 to M with velocity c, we have

$$S_0M = x_0 + \frac{v}{c} S_0M, \quad S_0M = \frac{x_0}{1-\beta}.$$

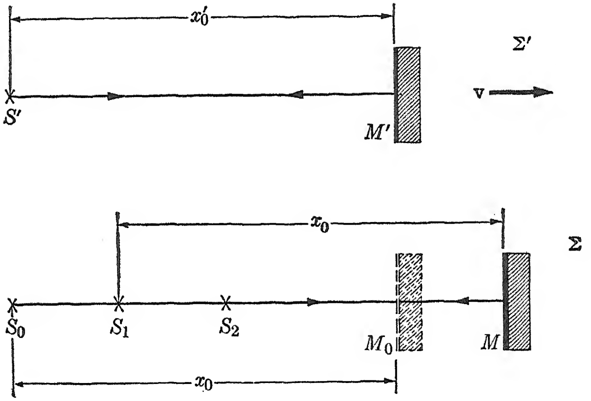


FIG. 15–3 Comparison of lengths parallel to the direction of motion.

Similarly, since the source has moved from S_1 to S_2 with velocity v while the light has traveled from M to S_2 with velocity c,

$$MS_2 = x_0 - \frac{v}{c} MS_2, \quad MS_2 = \frac{x_0}{1+\beta}.$$

Hence,

$$\Delta t = \frac{S_0M + MS_2}{c} = \frac{2x_0}{c(1-\beta^2)}, \tag{15–6}$$

and with the use of Eqs. (15–5) and (15–4), we obtain

$$x_0 = x_0'\sqrt{1-\beta^2}. \tag{15–7}$$

This relation, called the Lorentz contraction, is asymmetrical in x_0 and x_0', since it gives the relation between measurement of a proper length x_0' (at rest) in Σ' and an improper length x_0 (not at rest) in Σ. The length x_0 in Σ has been definable only by the assumption of the constancy of the velocity of light. Equation (15–7) is identical with the hypothesis of Section 14–3, but here it is accompanied by a time dilation quite foreign to the postulate of Lorentz and Fitzgerald.

EXPERIMENT IV. *The synchronization of clocks.* By appropriate design of the conceptual experiments II and III, we have been able to derive from the postulates of relativity the transformation of temporal and spatial intervals from proper to nonproper frames. The question of how two clocks, synchronized in frame Σ' but separated by a distance x_0', would appear to an observer in Σ requires further consideration.

Consider two clocks synchronized in Σ' and located a distance x_0' apart, as indicated in Fig. 15–4. Let there be a single clock in Σ which will re-

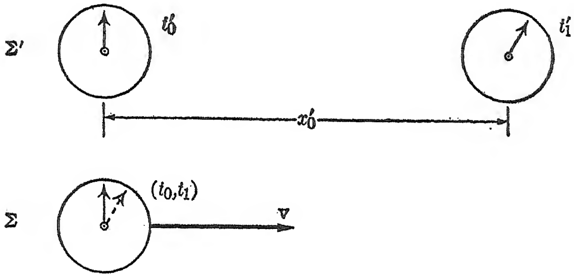


FIG. 15-4 Synchronization of clocks.

cord the times t_0 and t_1 when it passes the ends of x'_0. The corresponding times as recorded on the two clocks in Σ' are denoted by t'_0 and t'_1. Since $t_1 - t_0$ is a proper time interval in Σ, we can apply Eq. (15-4) in reverse, i.e.,

$$t'_1 - t'_0 = \frac{t_1 - t_0}{\sqrt{1 - \beta^2}}. \tag{15-8}$$

To the observer in Σ, the time intervals on both of the Σ' clocks individually would appear too large by the same ratio, but they will not be in time with each other. He would observe an error in synchronization of amount δ, that is,

$$t_1 - t_0 = \frac{t'_1 - t'_0 + \delta}{\sqrt{1 - \beta^2}}. \tag{15-9}$$

To determine δ, we need only note that in both frames the relative velocity must be the same, since the frames are equivalent. Hence

$$\begin{aligned} t'_1 - t'_0 &= x'_0/v, \\ t_1 - t_0 &= x_0/v. \end{aligned} \tag{15-10}$$

But x'_0 is a proper length in Σ', so that

$$x_0 = x'_0\sqrt{1 - \beta^2}. \tag{15-7}$$

By combining Eqs. (15-9), (15-10), and (15-7), we obtain

$$\delta = -\frac{x'_0\beta^2}{v}. \tag{15-11}$$

The meaning of the negative sign is that to an observer in Σ the second clock in Σ' indicates a time later than the first. A whole series of clocks, equally spaced and synchronized in Σ', would appear successively ahead by an amount δ to an observer in Σ as he went by with velocity v.

By means of these four conceptual experiments we have demonstrated that the postulates of special relativity lead to four kinematic relations:

I. Spatial intervals transverse to the direction of relative motion are invariant.

II. A time interval Δt between two events measured in a frame moving with velocity $\pm v$ relative to a frame in which the corresponding time interval $\Delta\tau$ is proper (i.e., the two events occur at one place) is given by

$$\Delta t = \frac{\Delta\tau}{\sqrt{1-\beta^2}}. \tag{15–12}$$

III. The length Δx of a rod measured in a frame moving with velocity $v_x = \pm v$ relative to a frame in which the rod is at rest and has the proper length ΔL is given by

$$\Delta x = \Delta L\sqrt{1-\beta^2}. \tag{15–13}$$

IV. Two clocks, synchronous and separated by a distance ΔL in a given frame, appear out of synchronism as observed from a frame moving with a relative velocity v to the clock frame by an amount

$$\delta = -\frac{\Delta L v}{c^2}. \tag{15–14}$$

15–3 The Lorentz transformation. The separate kinematic effects derived in the previous section from the fundamental postulates of relativity can be combined to give the general relations between the time and space coordinates of a particular event as observed from inertial frames in relative motion.

A point event at P moving with the Σ' frame, as indicated in Fig. 15–5, occurs at time t' and with coordinates x', y', z' in Σ'. Let us consider this same event as observed from a frame Σ in motion relative to Σ'. For simplicity, we may choose the x- and x'-axis as the direction of relative motion, and let the origins and the zero point of time be so chosen that at

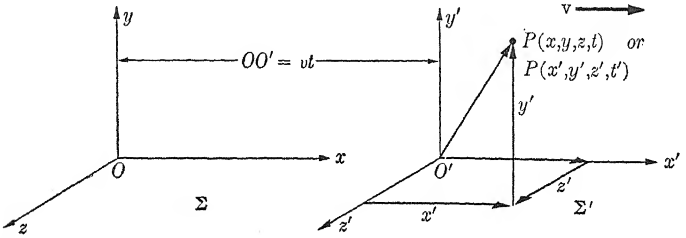


FIG. 15–5 Derivation of the Lorentz transformation equations for Σ and Σ′.

$t = t' = 0$ the two origins coincide. The time t or t' therefore means the time elapsed since the coincidence of origins as measured by an observer in Σ or Σ', respectively. We shall say that v is positive if the origin of Σ' moves along the positive x-direction in Σ.

To an observer in Σ, $OO' = vt$, but x', a proper length in Σ', is shortened by the Lorentz contraction. Hence

$$x = vt + x'\sqrt{1 - \beta^2} \tag{15–15}$$

or

$$x' = \frac{x - vt}{\sqrt{1 - \beta^2}}. \tag{15–16}$$

Again from the point of view of Σ, clocks located at P and O' (and synchronized in Σ') are out of synchronism by an amount

$$\Delta t' = \frac{x'v}{c^2}. \tag{15–17}$$

The clocks at O and O' were synchronous at $t = t' = 0$. Since that time, according to Σ, the Σ' clocks have been running slow, i.e., at a rate which must be dilated by $1/\sqrt{1 - \beta^2}$ to make it equal to the rate of the clock on Σ. Combining this effect with that of Eq. (15–17), we have

$$t = \frac{t' + x'v/c^2}{\sqrt{1 - \beta^2}}. \tag{15–18}$$

By the use of Eq. (15–16) this can be reduced to

$$t' = \frac{t - xv/c^2}{\sqrt{1 - \beta^2}}. \tag{15–19}$$

Equations (15–18) and (15–19) indicate that except for the sign of v, Σ and Σ' are equivalent, in agreement with the second relativity postulate. It can also be shown from Eqs. (15–18) and (15–16) that

$$x = \frac{x' + vt'}{\sqrt{1 - \beta^2}}, \tag{15–20}$$

in agreement with Eq. (15–16) except for the sign of v.

Equations (15–16), (15–20), (15–18), and (15–19), together with the consequence of Experiment I of Section 15–2 that lengths perpendicular to the motion are unaffected, constitute the general transformations we have sought, subject to the restrictions as to choice of origin and orientation of axes given above. For convenience, we may summarize what is known as the Lorentz transformation:

$$\begin{aligned} x' &= \frac{x - vt}{\sqrt{1-\beta^2}}, \\ y' &= y, \\ z' &= z, \\ t' &= \frac{t - xv/c^2}{\sqrt{1-\beta^2}}. \end{aligned} \tag{15–21}$$

It is easy to show algebraically that if the Lorentz transformation is valid,

$$x^2 + y^2 + z^2 - c^2t^2 = x'^2 + y'^2 + z'^2 - c^2t'^2. \tag{15–22}$$

This means that if a light signal is propagated in all directions with velocity c from O at $t = 0$, as observed in frame Σ, then a light signal is propagated from O' in all directions with velocity c at $t' = 0$, as observed in Σ'. The transformations are therefore in agreement with the first postulate, and resolve the apparent paradox mentioned at the beginning of the chapter.

The Galilean transformation equations, (14–1), do not satisfy Eq. (15–22). For two events, $(x_1,y_1,z_1;t_1)$ and $(x_2,y_2,z_2;t_2)$, Eqs. (14–1) yield

$$\begin{aligned} (x_1' - x_2')^2 + (y_1' - y_2')^2 + (z_1' - z_2')^2 \\ = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2, \\ (t_1' - t_2') = (t_1 - t_2), \end{aligned}$$

showing that in prerelativistic physics the spatial interval and the temporal interval between two events are *independently* invariant. In special relativity it is the *combined* space-time interval

$$(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 - c^2(t_1 - t_2)^2$$

which is invariant. In terms of the differential interval between two events the quantity

$$ds^2 = c^2\,dt^2 - dx^2 - dy^2 - dz^2 \tag{15–23}$$

is invariant under a Lorentz transformation.

The differential interval ds defined by Eq. (15–23) involves both space and time, but not symmetrically. If in any frame

$$dx^2 + dy^2 + dz^2 < c^2\,dt^2, \tag{15–24}$$

then an inertial frame can be found in which the spatial part of ds is zero, i.e., in which the two events occur at the same place. In that frame, $ds = c\,dt$ is just c times a proper time interval, as defined earlier. If the space-time interval satisfies the inequality (15–24), we say that the interval is "time-like," and ds/c represents the proper time interval between the events. Conversely, if in any frame

$$dx^2 + dy^2 + dz^2 > c^2\,dt^2, \tag{15–25}$$

then a frame can be found in which $dt = 0$; in that frame the two events are simultaneous and $i\,ds$ is their mutual distance. Hence, if ds satisfies the inequality (15–25), the interval is space-like, and $i\,ds$ represents the proper length of the increment. No Lorentz transformation with real β can invalidate either of the inequalities (15–24) or (15–25), and the time-like or space-like nature of an interval ds is an invariant.

It can be shown algebraically that two successive Lorentz transformations with velocity parameters β_1 and β_2 are equivalent to a single Lorentz transformation of parameter

$$\beta = \frac{\beta_1 + \beta_2}{1 + \beta_1\beta_2}. \tag{15–26}$$

The Lorentz transformations thus form a mathematical "group" with "commutative" properties.

It is possible to obtain the Lorentz transformation equations in several ways by simply using the demand that the interval ds of Eq. (15–23) be invariant and that the transformations be linear; the second condition arises from the fact that all points in space and time should have identical transformation characters if only inertial frames are considered. We must also demand, in accord with the relativity principle, that if $x = f(x',t',v)$, then $x' = f(x,t,-v)$. As an example of such a derivation, let us begin by assuming

$$kx = x' - vt', \quad kx' = x + vt, \tag{15–27}$$

where k is to be an even function of the velocity. By simple algebra,

$$x'^2 - x^2 = v^2(t^2 - t'^2)/(1 - k^2). \tag{15–28}$$

To make Eq. (15–28) agree with the invariant interval, Eq. (15–22), we must set

$$c^2(1 - k^2) = v^2$$

or

$$k = \sqrt{1 - \beta^2}, \tag{15–29}$$

giving the Lorentz transformation.

15–4 Geometric interpretations of the Lorentz transformation. The Lorentz transformation treats x and t as equivalent variables. It was suggested by Minkowski that ct be introduced simply as a fourth coordinate. Let us put

$$x^1 = x, \quad x^2 = y, \quad x^3 = z, \quad x^4 = ct, \tag{15–30}$$

as a set of variables in four-dimensional space. (Superscripts rather than subscripts are used here for a reason that will become clear in the following chapters.) The space-time interval ds is therefore given by

$$ds^2 = -(dx^1)^2 - (dx^2)^2 - (dx^3)^2 + (dx^4)^2. \tag{15–31}$$

The Lorentz transformation is thus, in a general sense, the set of linear transformations in four-dimensional space which leaves ds^2 invariant.

The interval ds can be written more symmetrically, although the result is physically less obvious, if we introduce

$$X^4 = ix^4, \quad dS = i\,ds. \tag{15–32}$$

Equation (15–23) becomes

$$dS^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2 + (dX^4)^2. \tag{15–33}$$

Since our choice of coordinates in the Lorentz transformation leaves dx^2 and dx^3 unaffected, it will suffice to consider the invariance of the two-dimensional interval

$$dS^2 = (dx^1)^2 + (dX^4)^2. \tag{15–34}$$

This interval is invariant to translations of the origin, and also to rotations of the coordinate axes in the x^1–X^4 plane. Let the coordinate axes be rotated through an angle θ, as shown in Fig. 15–6. It is clear that

$$\begin{aligned} x^{1\prime} &= x^1 \cos\theta + X^4 \sin\theta, \\ X^{4\prime} &= -x^1 \sin\theta + X^4 \cos\theta. \end{aligned} \tag{15–35}$$

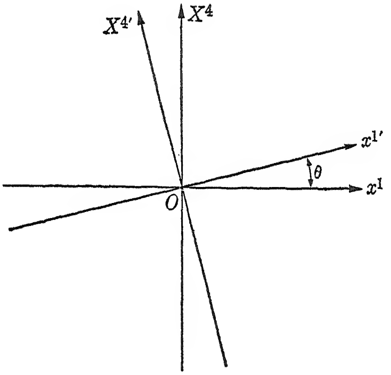


FIG. 15–6 Rotation of the complex coordinate axes through the angle θ.

Putting $\theta = i\varphi$, and transforming back to ordinary space and time variables by Eqs. (15–30) and (15–32), we obtain

$$\begin{aligned} x' &= x \cosh\varphi - ct \sinh\varphi, \\ ct' &= -x \sinh\varphi + ct \cosh\varphi. \end{aligned} \tag{15–36}$$

Equations (15–36) are identical with the Lorentz transformation equations if we put

$$\sinh\varphi = \frac{\beta}{\sqrt{1-\beta^2}}\,; \quad \cosh\varphi = \frac{1}{\sqrt{1-\beta^2}}\,; \quad \tanh\varphi = \beta. \qquad (15\text{–}37)$$

Thus the Lorentz transformation is simply a rotation in the four-dimensional space x^1, x^2, x^3, X^4. An "event" is therefore conveniently described by the four coordinates in such a space-time system, where temporal and spatial coordinates are entirely equivalent. Equation (15–26), which gives the β equivalent to two successive Lorentz transformations, simply corresponds to the addition formula for $\tanh\varphi$:

$$\tanh(\varphi_1 + \varphi_2) = \frac{\tanh\varphi_1 + \tan\varphi_2}{1 + \tanh\varphi_1 \tanh\varphi_2}\cdot \qquad (15\text{–}38)$$

The representation of the Lorentz transformation as a rotation in the four-dimensional space x, y, z, ict is a useful concept, but it is artificial from a physical viewpoint. Let us investigate the geometrical representation of the Lorentz transformation in the real four-dimensional space of x, y, z, ct. We shall plot only $x^1 = x$ and $x^4 = ct$, so as to permit representation in a plane. The trajectory of an event so plotted as a function of space and time is called a world line, and the diagram itself is called the Minkowski diagram. The world line of a ray of light *in vacuo* is the line $x^1 = x^4$, at 45° to each of the original axes in Fig. 15–7. Under a Lorentz transformation, these axes will transform into $x^{1\prime}$ and $x^{4\prime}$ by Eqs. (15–36), but the world line of the light ray is unchanged. Figure 15–7 shows that simultaneity is a relative concept: all events located on the $x^{1\prime}$ axis are simultaneous in Σ', but not in Σ. Thus event P' is simultaneous with event O to an observer at rest in Σ', but occurs later, at time $t = x^4/c = PP'/c$, to an observer at rest in Σ.

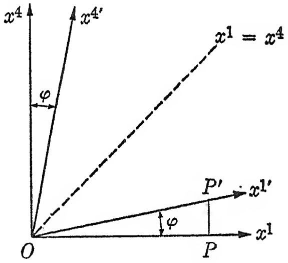


FIG. 15–7 Lorentz transformed axes in a Minkowski diagram, with PP' illustrating the relativity of simultaneity.

Considerable care is necessary in the general interpretation of the Minkowski diagram. In contrast to the diagram of Fig. 15–6, where we artificially produced a "Euclidian" geometry by the imaginary transformation Eqs. (15–32), the intervals cannot be measured by the sum of the squares of the coordinate intervals. A substitute for distance measurement in this real space can be obtained by noting that the family of hyperbolas

$$(x^1)^2 - (x^4)^2 = (x^{1\prime})^2 - (x^{4\prime})^2 = \text{constant} \qquad (15\text{–}39)$$

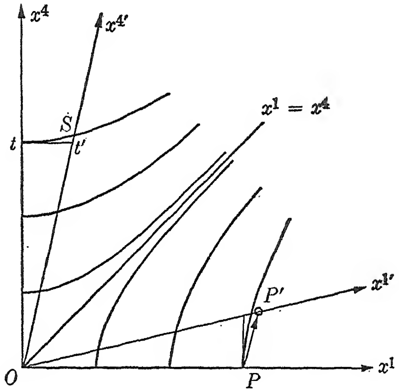


FIG. 15–8 Minkowski diagram showing Lorentz contraction and time dilation.

lays out a convenient net which permits comparison of the various quantities involved. Such a net is indicated in Fig. 15–8.

Let us see how the various phenomena of relativistic kinematics are interpreted on this diagram. The Lorentz contraction, Eq. (15–7), gives the transformation of a proper length $x^{1\prime}$ in Σ' to the Σ frame. Consider a rod OP' at rest in Σ'. The world line of the end point P' is moving parallel to the $x^{4\prime}(ct')$-axis, since it is at rest (proper) in Σ'. Similarly, the point O is moving along the $x^{4\prime}$ axis. The rod is measured in Σ when the origins coincide and when the end points are simultaneous in Σ, i.e., along the x^1-axis. The length of the rod in Σ is thus the length OP. In comparing OP and OP', we must be careful to refer the measurements to the hyperbolic grid. It is easily seen that the hyperbola

$$(x^1)^2 - (x^4)^2 = OP^2 = (x^{1\prime})^2 - (x^{4\prime})^2$$

crosses $Ox^{1\prime}$ between O and P', and hence $OP < OP'$, in accordance with the Lorentz contraction of proper lengths observed from a moving frame.

In order to consider the time dilation, let us take a single clock at rest in Σ' at $x' = t' = x = t = 0$. As time progresses, the time interval relative to $t' = t = 0$ will be represented by a world line moving along the $x^{4\prime}$-axis and, in Σ', is thus just Ot'. O and t' are not at the same spatial point in Σ, and the point considered simultaneous with t' will be at t; there tt' is parallel to the x^1-axis. But in the hyperbolic "metric," $Ot = OS$, and $Ot > Ot'$. Hence an observer at rest in Σ will observe a longer elapsed time than the proper time interval measured in Σ', in agreement with the time dilation of Eq. (15–4).

The Minkowski diagram shows the symmetry between the Σ and Σ' frames despite the apparent asymmetry of the time dilation and Lorentz contraction. In our examples, Σ' was taken as the proper frame for both spatial and temporal intervals. If we had chosen Σ as the proper frame,

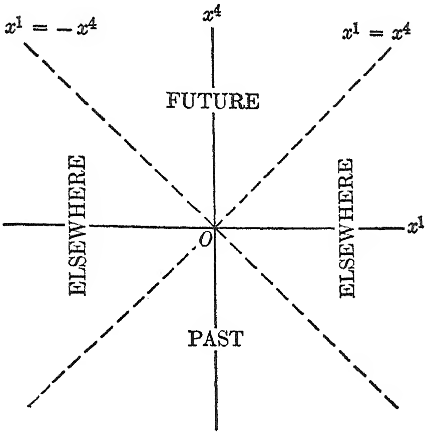


FIG. 15-9 Minkowski diagram showing the division of space-time by the light cone into past, future and elsewhere for an observer at the origin.

PP' would have been parallel to Ot, and tt' parallel to OP', which would have reversed the contraction and dilation relations.

A Lorentz transformation with $\beta > 1$ is complex and thus physically impossible. (We shall re-examine the significance of this statement later.) On the Minkowski diagram, this means that neither of the primed axes can pass the line $x^1 = -x^4$, which in four-dimensional space is a cone, called the light cone. Hence a time-like interval with $ds^2 = (dx^4)^2 - (dx^1)^2 > 0$ cannot become space-like in any frame, but can become purely temporal if referred to the proper frame. For the same reason, a space-like interval $ds^2 < 0$ cannot become time-like in any frame, but can become purely spatial if referred to a proper frame.

The light cone $x^1 = \pm x^4$ thus divides the Minkowski space into regions, as indicated in Fig. 15-9, that have invariant significance for an observer at the origin, i.e., at a given time and place. The temporal region labeled "past" represents events whose temporal interval relative to the origin is negative from any inertial frame. Similarly, positive temporal intervals are confined to the region labeled "future," regardless of the frame of the observer at O. Events in the region called "elsewhere" are spatially separated from the observer, but can be transformed into one another by spatial rotations and translations.

15-5 Transformation equations for velocity. The transformation equations for the velocity of a moving point can be found by taking the derivatives of the Lorentz transformation equations with respect to t and t'. We shall continue to use the symbols v and $\beta = v/c$ to denote the velocity of frame Σ' relative to frame Σ, evaluated in Σ, and introduce $u_x = dx/dt$, $u'_x = dx'/dt'$, etc., to denote velocities in a given frame. The Lorentz transformation from Σ to Σ' has been found to be

$$x' = \frac{x - vt}{\sqrt{1 - \beta^2}},$$
$$y' = y, \quad z' = z, \tag{15-40}$$
$$t' = \frac{t - \beta x/c}{\sqrt{1 - \beta^2}}.$$

Differentiating Eqs. (15–40) with respect to t', we obtain

$$\frac{dx'}{dt'} = u_x' = \frac{u_x - v}{\sqrt{1 - \beta^2}} \frac{dt}{dt'}, \tag{15-41}$$

$$\frac{dy'}{dt'} = \frac{dy}{dt}\frac{dt}{dt'}, \quad \frac{dz'}{dt'} = \frac{dz}{dt}\frac{dt}{dt'}. \tag{15-42}$$

Also,

$$\frac{dt'}{dt} = \frac{1 - \beta u_x/c}{\sqrt{1 - \beta^2}}. \tag{15-43}$$

Equation (15–43) enables us to write Eqs. (15–41) and (15–42) in terms of the unprimed variables:

$$u_x' = \frac{u_x - v}{1 - u_x v/c^2}, \tag{15-44}$$

$$u_y' = \frac{\sqrt{1 - \beta^2}}{1 - u_x v/c^2} u_y, \quad u_z = \frac{\sqrt{1 - \beta^2}}{1 - u_x v/c^2} u_z. \tag{15-45}$$

Equation (15–44), the "longitudinal velocity addition formula," is in agreement with Eq. (15–26) for successive Lorentz transformations, since u_x may represent the motion of the origin of another Lorentz frame relative to the Σ' frame.

Equation (15–43), the relation between time intervals dt and dt', has some interesting consequences. If it were possible to make $(dt'/dt) < 0$ by a suitable choice of u, then the temporal sequence of two events would be reversed between the two frames under consideration. This would be a logical contradiction if (a) the two events represent cause and effect, and (b) the *sense* of time has an invariant significance. With the assumption of (b), made at the beginning of this chapter, we are forced to conclude that, in order for the sequence between cause and effect to be preserved in all frames, in *any* particular frame

$$u \leq c, \tag{15-46}$$

where u represents the velocity of propagation of any event which can connect cause and effect. Obviously, phase velocities, or velocities of geometrical significance only, are not affected by this restriction. The restriction does apply to the relative velocities v of possible inertial frames, so that it is not necessary to consider the kinematical significance of the Lorentz transformation when $\sqrt{1-\beta^2}$ is complex.

The velocity transformation equations apply very simply to the velocity of light in a moving medium and to stellar aberration.

Let a homogeneous isotropic medium of index of refraction n move with velocity v along the positive x-axis with respect to an observer at rest in Σ. The velocity of light with respect to the frame in which the medium is at rest is $u' = c/n$. With respect to an observer at rest in Σ, the velocity is u, where u is obtained from Eq. (15–44) by reversing the sign of v,

$$u = \frac{c/n + v}{1 + v/nc} \simeq \frac{c}{n} + v(1 - 1/n^2) \tag{15–47}$$

to the first order in v/c. This is in agreement with the experimental facts and the classical electrodynamic result derived in Section 11–4. In the classical derivation of this equation, however, a relatively complicated mechanism was involved: the effect was attributed to reradiation from the moving secondary radiators in the fluid. Here Eq. (15–47) has been derived without any detailed information concerning the mechanism. We shall frequently meet situations in which an end result is demanded by relativistic considerations, but where the physical mechanism of attaining the result is far from obvious.

The aberration of distant stars can also be derived from the velocity transformation equations. Let a ray of starlight approach the earth in a direction perpendicular to the earth's velocity, as indicated in Fig. 15–10. In the frame of the star the process is simply the emission of light with velocity components $u_y = -c$, $u_x = 0$. In the earth's frame, the velocity components as given by Eqs. (15–44) and (15–45) are

$$u'_y = -c\sqrt{1 - v^2/c^2}, \quad u'_x = -v$$

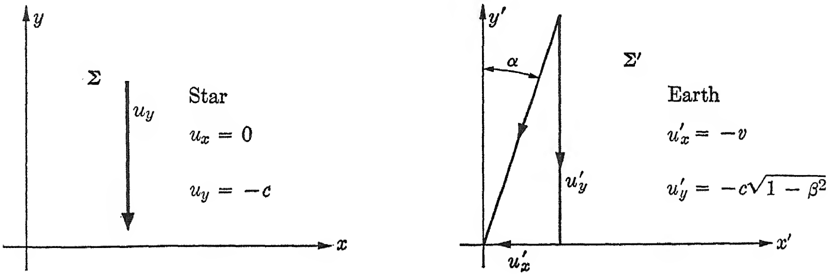


FIG. 15–10 Relativistic explanation of aberration of starlight.

Therefore the angle of incidence to the normal is given by

$$\begin{aligned} \tan\theta' &= \frac{u'_x}{u'_y} = \frac{\beta}{\sqrt{1-\beta^2}}, \\ \sin\theta' &= \beta. \end{aligned} \tag{15-48}$$

We have seen in Chapter 14 that a mechanical emission picture or the assumption of a stationary ether gives $\tan\theta' = \beta$, which is, in practice, indistinguishable from Eq. (15-48).

SUGGESTED REFERENCES

There are many excellent introductions to the concepts of special relativity and the Lorentz transformation. The list below includes treatments most familiar to the present authors and those likely to be most easily accessible.

R. BECKER, *Theorie der Elektrizitat*, Band II, especially pp. 266–282. The Minkowski geometry is particularly well presented.

P. G. BERGMANN, *An Introduction to the Theory of Relativity*. Chapter IV introduces the Lorentz transformation in a very lucid way.

R. C. TOLMAN, *Relativity, Thermodynamics, and Cosmology*, through p. 34, or, if it is available, *The Theory of the Relativity of Motion*, which provides an excellent introduction to the subject.

A. EINSTEIN AND H. MINKOWSKI in *The Principle of Relativity*. Original papers, carefully translated.

A. EINSTEIN, *The Meaning of Relativity*. This is not a textbook nor one designed to be read hurriedly. It will repay careful study.

C. MØLLER, *The Theory of Relativity*, Chapters I and II. This is a recent book to which we shall refer in later chapters.

Of the many more general textbooks which contain clear and competent introductions to special relativity, we may mention

G. JOOS, *Theoretical Physics*, Chapter X.

R. B. LINDSAY AND H. MARGENAU, *Foundations of Physics*, Chapter 7.

EXERCISES

1. Does the choice of relative velocity along the x-axis and coincidence of origins at $t = t' = 0$ put any essential limitations on the Lorentz transformation as given by Eqs. (15-21)? How, in detail, would you handle an arbitrarily directed relative velocity and an arbitrary pair of origins in the coordinate frames in making a Lorentz transformation?

2. With relation to the Minkowski diagram of Fig. 15-8, interpret a proper length and a proper time in Σ as observed in the primed system Σ'.

3. Show that the transformation of $\dot{u}^2$ from a proper frame to a general frame is

$$\dot{u}^2 \rightarrow [\dot{u}^2 - (\dot{\mathbf{u}} \times \boldsymbol{\beta})^2]\gamma^6,$$

where $\gamma = 1/\sqrt{1-\beta^2}$.

CHAPTER 16

COVARIANCE AND RELATIVISTIC MECHANICS

In Chapter 15 we have investigated the bearing of the principles of special relativity on the laws of kinematics. If the principles are valid they must apply to all fields of physics; by no experiment of any kind should it be possible to detect a preferred inertial frame. The bearing of the principles of relativity on other fields could be introduced by designing appropriate *Gedanken Experimente,* as we did in kinematics. Or we could attempt to obtain transformation relations for physical quantities by applying the Lorentz transformation to the time and space coordinates of the pertinent prerelativistic equations and then trying to deduce transformation relations for the remaining quantities. Both of these approaches are useful. The thought experiment approach remains closest to physical concepts, and we shall make use of it in discussing collisions. The direct transformation process is frequently tedious, although it was the method used by Einstein in his original work to deduce the transformation equations for the electromagnetic fields and to show that Maxwell's equations are in agreement with relativistic principles.

There is a third method which is by far the most powerful and efficient in extending relativity. This approach is to rewrite the equations in a *form* which explicitly makes evident how the quantities would behave under a change to a different inertial frame. If an equation has a form which is invariant to a change in inertial frame, then an experiment based on this equation obviously could not give a result depending on the particular frame of reference. The equation then describes a phenomenon which would be in agreement with the principles of special relativity. An equation written in such a way that its form is independent of the choice of inertial frame is said to be *Lorentz covariant.*

We shall investigate this method sufficiently to be able to deduce the basic relativistic relations in mechanics and electrodynamics.

16–1 The Lorentz transformation of a four-vector. The Lorentz transformation can be written as a linear transformation of the interval components from the origin to point x^j, $(x^1,x^2,x^3,x^4 = ct)$, in a four-dimensional space. Making use of the summation convention for repeated indices, we may write

$$x^{i'} = Q^i_j x^j, \tag{16–1}$$

where Q^i_j is given by the matrix

$$Q^i_j = \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix}. \tag{16–2}$$

Here $\beta = v/c$ as usual, and for convenience in writing,

$$\gamma = \frac{1}{\sqrt{1-\beta^2}}. \tag{16–3}$$

We have already noted that if the components of the vector x^j transform according to Eq. (16–1), then an experiment involving x^j cannot yield a preferred frame. If, therefore, any physical relation is written as a vector equation in four-space where the vector components transform in accordance with Eq. (16–1), then such an equation is said to be written in Lorentz covariant form.

If we solve Eq. (16–1) for x^j, we obtain

$$x^j = (Q^i_j)^{-1} x^{i\prime}, \tag{16–4}$$

where

$$(Q^i_j)^{-1} = \begin{pmatrix} \gamma & 0 & 0 & +\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ +\beta\gamma & 0 & 0 & \gamma \end{pmatrix} \tag{16–5}$$

is the inverse matrix of Eq. (16–2), i.e., the matrix of the transformation corresponding to relative motion of the frames with opposite velocity. If a quantity with four components A_j transforms as the reverse transformation of the x^j, i.e., as

$$A'_i = (Q^i_j)^{-1} A_j, \tag{16–6}$$

then a relation equating components of the type A_j is also Lorentz covariant. To restate these two cases: any quantities A_j or B^j are Lorentz covariant if under change of inertial frame they transform as Eq. (16–6) or as

$$B^{i\prime} = Q^i_j B^j, \tag{16–7}$$

respectively. A_j is called a covariant four-vector and B^j is called a contravariant four-vector. (It is unfortunate that the word "covariant" is used in two different ways, but this usage is a matter of accepted convention.) In the language of matrices Q^i_j is related to the inverse transformation by the equation

$$(Q^k_j)^{-1} Q^i_j = \delta^i_k, \tag{16–8}$$

where

$$\delta^i_k = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{16-9}$$

is a matrix representing identity.

We are not here concerned with general tensor analysis *per se*, nor with the general transformation of coordinates. In the Lorentz transformation of special relativity the transformation coefficients Q^i_j are constants and thus independent of the coordinates. In most of the references this transformation with constant coefficients is derived as a special case of the general transformation of coordinates in which the coefficients themselves are functions of the coordinates. In the general analog of Eq. (16-1) a differential coordinate element dx^j transforms like

$$dx^{i'} = \left(\frac{\partial x^{i'}}{\partial x^j}\right) dx^j. \tag{16-10}$$

Covariant and contravariant vector components transform as

$$A'_i = \left(\frac{\partial x^j}{\partial x^{i'}}\right) A_j, \quad B^{i'} = \left(\frac{\partial x^{i'}}{\partial x^j}\right) B^j, \tag{16-11}$$

respectively. In the general theory of relativity it is necessary to use Eqs. (16-10) and (16-11), since the transformation coefficients may be functions of the coordinates. The distinction between covariant and contravariant entities takes on added significance in that case. In special relativity it is clear from Eq. (16-5) that the covariant vector A_j simply transforms like $-x$, $-y$, $-z$, ct. Even when the transformation coefficients are constants, however, the algorism of general tensor analysis is extremely useful, and we shall follow the standard forms.

16-2 Some tensor relations useful in special relativity. For our purpose, we may define as tensors all quantities which themselves maintain definite transformation properties when the coordinates undergo a Lorentz transformation. The simplest is an invariant, or scalar, which can be specified by a single number. We have seen, for example, that $ds^2 = c^2\,dt^2 - dx^2 - dy^2 - dz^2$ is such an invariant. A quantity of this kind is called a tensor of rank zero.

The contravariant and covariant four-vectors of Section 16-1 are tensors of rank one. A contravariant tensor of rank two is a collection of sixteen quantities which transforms like

$$T^{kl'} = Q^k_i Q^l_j T^{ij}, \tag{16-12}$$

i.e., like the product of two contravariant vectors. Similarly, a covariant tensor of second rank is a set of sixteen components which transform like

$$T'_{kl} = (Q_i^k)^{-1}(Q_j^l)^{-1}T_{ij}, \tag{16–13}$$

i.e., like the product of two covariant vectors. A mixed tensor of second rank transforms like the product of a contravariant and a covariant vector,

$$T_k^{l\,\prime} = Q_j^l(Q_i^k)^{-1}T_i^j. \tag{16–14}$$

Thus the pattern becomes clear: any quantity of the type $T^{l_1l_2\cdots l_n}_{k_1k_2k_3\cdots k_m}$ which transforms like

$$T^{l_1l_2\cdots l_n\,\prime}_{k_1k_2k_3\cdots k_m} = Q^{l_1}_{i_1}Q^{l_2}_{i_2}\cdots Q^{l_n}_{i_n}(Q^{k_1}_{j_1})^{-1}(Q^{k_2}_{j_2})^{-1}\cdots(Q^{k_m}_{j_m})^{-1}T^{i_1i_2\cdots i_n}_{j_1j_2\cdots j_m}, \tag{16–15}$$

i.e., simply like products of covariant and contravariant four-vector components, is called a tensor of rank $(m + n)$. It follows that the product of a tensor of rank m and one of rank n is a tensor of rank $m + n$. We shall rarely be concerned with tensors of higher than second rank in our physical applications.

Since the transformations involved are linear, the sum or difference of any two tensors of equal rank is a tensor of the same rank, and equations can be made only between quantities of the same kind. Such tensor equations will be Lorentz covariant. In order to extend the principles of relativity by rewriting other laws of physics in the form of tensor equations, we shall need some of the formal rules of tensor manipulation.

1. When a mixed tensor has a contravariant index which is the same as a covariant index the implied summation reduces the rank of the tensor by two. This process is called *contraction*. The contraction of a tensor T^i_j, namely, T^i_i, is an invariant. We shall see that if we contract the vector gradient (a tensor of rank 2) we obtain the divergence (a tensor of rank 0).

2. We have seen that the "line element" in special relativity, $ds^2 = -dx^2 - dy^2 - dz^2 + c^2\,dt^2$, is a scalar invariant. ds^2 can be written in the form

$$ds^2 = g_{ij}\,dx^i\,dx^j, \tag{16–16}$$

where

$$g_{ij} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & +1 \end{pmatrix} \tag{16–17}$$

is called the "metric tensor" corresponding to the line element. (In general tensor analysis g_{ij} is a function of the coordinates.) It can easily be proved by transformation, using Eq. (16–5), that g_{ij} is actually a tensor.

3. By means of the relation

$$B_i = g_{ij}B^j \tag{16–18}$$

each member of a set of covariant components B_i can be associated with a contravariant tensor component. For the special form of g_{ij} given by Eq. (16–17) this process will simply reverse the sign of the first three components. The covariant components dx_i corresponding to the basic contravariant interval dx^i are such that

$$dx^i\, dx_i = ds^2. \tag{16–19}$$

4. With reference to Eqs. (16–11), it can be seen by writing $\partial/\partial x^{i\prime} = (\partial x^j/\partial x^{i\prime})\, \partial/\partial x^j$ that the derivative $\partial/\partial x^i$ operating on a tensor transforms like an additional covariant tensor factor. (This is not true if the g_{ij} are functions of the coordinates.) The operation of the "gradient" thus increases the rank of the tensor by one (covariant index): if T^0 is a scalar, its gradient $\partial T^0/\partial x^i$ is a covariant vector, and the increment of the scalar can be written as a tensor relation by contraction,

$$dT^0 = \frac{\partial T^0}{\partial x^i}\, dx^i. \tag{16–20}$$

Similarly, if T^k is a contravariant vector, $\partial T^k/\partial x^i$ is a mixed tensor of the second rank, while $\partial T^i/\partial x^i$, the four-divergence of T^i, is a scalar invariant.

5. If a four-vector obeys the relation

$$\frac{\partial T^i}{\partial x^i} = 0, \tag{16–21}$$

and if the components of T^i are different from zero only in a finite spatial region, then the integral over three-dimensional space,

$$I = \int T^4\, dv, \tag{16–22}$$

is an invariant. Let us prove this theorem. The four-dimensional analog of Gauss's theorem states that

$$\int \frac{\partial T^i}{\partial x^i}\, d^4x = \int T^i\, dS_i, \tag{16–23}$$

where $d^4x = dx^1\, dx^2\, dx^3\, dx^4 = dv\, dx^4$, and dS_i is an element of three-dimensional "surface" normal to T^i in four-space. The region over which the integration of Eq. (16–23) is to be performed is indicated in Fig. 16–1, where the surfaces (A) and (C) are chosen so that the spatial components of T^i vanish on (A) and (C). This can be done, for it was assumed that the region where the spatial components of T^i do not vanish is finite.

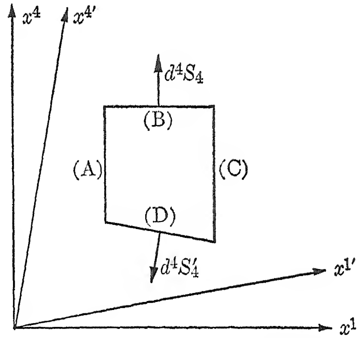


FIG. 16–1 Showing region of integration for proving the theorem of Eq. (16–22).

(B) is chosen normal to the x^4-axis, while (D) is chosen normal to the $x^{4\prime}$-axis. It then follows from Eqs. (16–21) and (16–23) that

$$\int T^4\,dS_4 = \int T^{4\prime}\,dS_4'.$$

Since $dS_4 = dv$ and $dS_4' = dv'$, the ordinary three-dimensional volume element, this means that

$$I = \int T^4\,dv \tag{16–22}$$

is an invariant under a Lorentz transformation. Similarly, it can be proved that if a tensor of rank two satisfies the relation

$$\frac{\partial T^{ij}}{\partial x^i} = 0, \tag{16–24}$$

then

$$B^j = \int T^{4j}\,dv \tag{16–25}$$

is a four-vector. Equations (16–21) and (16–24) will later be recognized as "conservation laws."

6. A tensor possesses symmetry with respect to two contravariant indices, i and j, or two covariant indices i and j, if the interchange of i and j leaves all the components of the tensor unchanged. A tensor is said to be antisymmetric, or *skew-symmetric*, if such an interchange of i and j simply changes the sign of all components. All symmetry properties are invariant under a transformation of coordinates.

These relations will suffice for our covariant description of particle mechanics, although we have refrained from anticipating all the problems involved in the covariant description of electromagnetic fields. Before attempting the extension of relativistic principles to mechanics, let us note that a covariant relation in physics can be generated in three ways:

(a) The relation is known in a special inertial frame, such as the proper frame where the system under consideration is at rest. If it is possible to express this law in the form of a tensor equation which reduces to the correct relation in the special frame, then this tensor equation has general significance.

(b) A known tensor relation is converted into a new tensor equation by a covariant tensor operation. The simplest example is multiplication by an invariant, but we have also noted tensor multiplication, contraction, and covariant differentiation as covariant tensor operations.

(c) An equation is obtained from a relation valid in a special frame by transforming those quantities whose transformation properties are known, and deducing the correct transformations for the remaining quantities. These quantities are then expressible in tensor form. This process has been referred to at the beginning of the chapter as that of direct transformation and it is usually very tedious.

16–3 The conservation of momentum. Let us use the above considerations to formulate the law of conservation of momentum for mass points. The formal relations so derived will then be compared with the result of a "thought experiment" involving an inelastic collision as observed from various frames.

In prerelativistic physics, a particle of mass m which is moving with velocity $\mathbf{u}$ possesses linear momentum $\mathbf{p}$, where

$$\mathbf{p} = m\mathbf{u}. \tag{16–26}$$

The vector $\mathbf{u}$ does not constitute the first three components of a four-vector, for $\mathbf{u} = d\mathbf{r}/dt$, and while $d\mathbf{r}$ represents three components of the four-vector $dx^i = (d\mathbf{r}, c\,dt)$, dt is not an invariant. [If the spatial components of a four-vector conform to a standard three-dimensional vector, we shall, in enumerating the components of the four-vector, use regular vector notation, as in $x^i = (\mathbf{r}, ct)$.] However, if we divide the contravariant vector dx^i by the invariant line element ds, we do obtain a four-vector

$$u^i = \frac{dx^i}{ds}, \tag{16–27}$$

known as the four-velocity. ds is the proper time interval multiplied by c, or

$$ds = c\,dt\sqrt{1 - \frac{u^2}{c^2}}, \tag{16–28}$$

and therefore the components of u^i are given by

$$u^i = \left(\frac{\mathbf{u}}{c\sqrt{1 - u^2/c^2}}, \frac{1}{\sqrt{1 - u^2/c^2}}\right). \tag{16–29}$$

We note that u^i so defined is dimensionless, instead of having the dimensions of ordinary velocity.*

---

* In constructing four-vectors or tensors we must make sure that the various members of the matrix representing the tensor have the same dimensions; this is done by using appropriate powers of c. Just how this is done is a matter of convention, and several texts differ from our convention by powers of c. For example, u^i may be consistently defined as $u^i = (\mathbf{u}/\sqrt{1 - u^2/c^2}, c/\sqrt{1 - u^2/c^2})$, which does give u^i the dimension of velocity. The convention of choosing units of length and time to make $c = 1$ is also frequently used, with resulting simplification.

A covariant expression corresponding to linear momentum can be generated by multiplying Eq. (16–27) by an invariant quantity m_0c^2 which is assumed to be characteristic of the particle. Here m_0 is the rest mass, i.e., the mass measured in the proper frame of the particle. The four-momentum thus defined is

$$p^i = m_0 \frac{dx^i}{ds} c^2. \tag{16–30}$$

Defined in this way, p^i is Lorentz covariant, although we note that its dimensions are those of energy. If we want to assure that the law of conservation of momentum shall be preserved in the framework of relativity for two particles interacting at a point, then in order that this law be independent of the choice of inertial frame it must take the form

$$p_1^i + p_2^i = \text{constant} \tag{16–31}$$

before and after the collision. Equation (16–31) replaces the classical law

$$m_1u_1 + m_2u_2 = \text{constant}, \tag{16–32}$$

in which it is assumed that m_1 and m_2 are constant.

This discussion is restricted to direct interaction between particles, rather than "interaction at a distance," for two reasons. The total momentum of two separated particles at a "given" time has no meaning in relativity, and furthermore all interactions are propagated with finite velocity. Therefore exchange of momentum between separated particles has meaning only if each particle conserves momentum with a field acting on it, or if the interaction is carried by a particle interacting in succession with the two mass points. Strictly speaking, each of the mass points considered here must have zero extension in order that the discussion be rigorous.

The components of the four-momentum defined by Eq. (16–30) are

$$p^i = \left(\frac{m_0\mathbf{u}c}{\sqrt{1 - u^2/c^2}}, \quad \frac{m_0c^2}{\sqrt{1 - u^2/c^2}}\right), \tag{16–33}$$

where $\mathbf{u}$ is the ordinary velocity in a given frame. The first three components have the form

$$c\mathbf{p} = cm\mathbf{u}, \tag{16–34}$$

where

$$m = \frac{m_0}{\sqrt{1 - u^2/c^2}}. \tag{16–35}$$

Hence, if we require that the law of conservation of momentum be maintained in its classical form, Eq. (16–32), and that it be Lorentz covariant, the mass of a particle is no longer an invariant but will depend on its velocity in the particular reference frame. The "variation of mass with velocity" is an immediate consequence of formulating the law of conservation of momentum in a covariant manner.

16-4 Relation of energy to momentum and to mass. The relativistic form of the law of conservation of momentum implies not only the conservation of the three spatial components of p^i but also the conservation of the fourth component,

$$p^4 = \frac{m_0c^2}{\sqrt{1 - u^2/c^2}} = mc^2. \tag{16-36}$$

Let us investigate the physical significance of this quantity. The time rate of change of p^4 in a given frame is

$$\frac{dp^4}{dt} = \frac{d}{dt}\left(\frac{m_0c^2}{\sqrt{1 - u^2/c^2}}\right) = \mathbf{u} \cdot \frac{d}{dt}\left(\frac{m_0\mathbf{u}}{\sqrt{1 - u^2/c^2}}\right), \tag{16-37}$$

or

$$\frac{dp^4}{dt} = \mathbf{u} \cdot \frac{d\mathbf{p}}{dt}. \tag{16-38}$$

If we continue to measure the (three-dimensional) force by the time rate of change of momentum

$$\mathbf{F} = \frac{d\mathbf{p}}{dt}, \tag{16-39}$$

then

$$\mathbf{F} \cdot \mathbf{u} = \mathbf{u} \cdot \frac{d\mathbf{p}}{dt} \tag{16-40}$$

represents the rate at which work is being done in a particular system. Hence if the law of conservation of energy is to hold in a particular frame, and if E denotes the energy in that frame,

$$\frac{dp^4}{dt} = \frac{dE}{dt}; \quad p^4 = E + \text{constant}. \tag{16-41}$$

Since energy manifests itself only when energy changes occur, we lose no physical significance if we put

$$E = mc^2 = \frac{m_0c^2}{\sqrt{1 - u^2/c^2}}. \tag{16-42}$$

We are thus led to conclude that energy as measured by "work content" and mass as measured by the momentum for a given velocity are interchangeable concepts; when one exists so does the other. Neither mass nor energy is an invariant; the magnitude of both depends on the frame of the observer by the relation of Eq. (16-42). We have shown that the change in mc^2 corresponds to work done by mechanical forces; that it corresponds generally to change in energy under any mechanism that might be involved implies an additional assumption whose justification rests with experiment. Experience in other fields of physics, particularly in nuclear physics where the fractional mass changes become very large, certainly proves beyond any reasonable doubt that Eq. (16–42) is valid in this more general interpretation.

For small velocities, E reduces to the classical kinetic energy plus the "rest energy" m_0c^2. By expansion of Eq. (16–42) for small values of u/c, we obtain

$$\begin{aligned} E &= m_0c^2\left[1 + \frac{1}{2}\left(\frac{u}{c}\right)^2 + \frac{3}{8}\left(\frac{u}{c}\right)^4 + \cdots\right] \\ &= m_0c^2 + \frac{1}{2}m_0u^2 + \frac{3}{8}m_0\frac{u^4}{c^2} + \cdots. \end{aligned} \tag{16–43}$$

In general, relativistic mechanics reduces to Newtonian mechanics in the limit of small velocities, as indeed it must.

Relativistically, the conservation of energy and the conservation of momentum are not independent principles; one demands the other for a covariant formulation. The invariant related to the energy-momentum four-vector, $p^i = (c\mathbf{p},E)$ in the same way that ds is related to dx^i may be easily ascertained:

$$p^ip_i = E^2 - c^2p^2 = (m_0c^2)^2. \tag{16–44}$$

This relation between energy and momentum is valid in any frame, and in a proper frame with $\mathbf{p} = 0$ we have simply $E = m_0c^2$. In other words, our assumption that m_0 is an invariant characteristic of a particular particle is vindicated, and by means of it we have achieved a covariant formulation of the conservation laws.

The energy-momentum vector p^i transforms like any other contravariant vector. In accordance with Eq. (16–2), we have

$$cp'_x = \frac{cp_x - \beta E}{\sqrt{1 - \beta^2}}, \tag{16–45}$$

$$p'_y = p_y; \quad p'_z = p_z, \tag{16–46}$$

$$E' = \frac{E - \beta cp_x}{\sqrt{1 - \beta^2}}, \tag{16–47}$$

where $E = p^4$. From these transformation relations, it can be seen that any transfer of energy implies transfer of mass and therefore momentum.

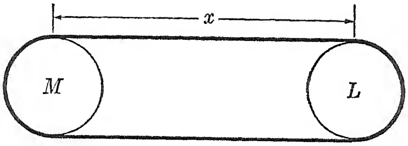


FIG. 16–2 Motor M driving load L at distance x.

If E is the energy of a system in its proper frame (zero momentum), then from the primed system we shall observe a momentum

$$p'_x = \frac{-\beta E}{\sqrt{1-\beta^2}} = \frac{-v}{c^2} E' \tag{16–48}$$

According to Eq. (16–48), we must associate momentum with any agency that transmits energy, whatever the form of the energy. Consider a motor M driving a load L at a distance x, as shown in Fig. 16–2. If the motor transfers energy at the rate dE/dt to the load, the mass of L increases correspondingly. Mass is being transmitted over a distance x at a rate of $(1/c^2)\, dE/dt$, and the system thus has a momentum

$$p = \frac{x}{c^2}\frac{dE}{dt}.$$

In general, if energy is being absorbed by a body at a given rate the momentum of the body increases, and to conserve over-all momentum we must associate a momentum density per unit volume, $\mathbf{g}$, with any agent that transmits energy at the rate $\mathbf{N}$ per unit area in a given direction, where

$$\mathbf{g} = \frac{\mathbf{N}}{c^2}. \tag{16–49}$$

This relation is in agreement with our discussion of radiation pressure in Chapter 11, and with our considerations of momentum balance in Section 10–6. When we considered electromagnetic radiation, which represents an energy flow $\mathbf{N}$ incident upon an observer, we were forced to attribute a momentum density as given by Eq. (16–49) to the electromagnetic field. This was really because we refused to accept the ether as capable of sustaining a volume force. The failure to detect an ether by the experiments outlined in Chapter 14 strengthens the conviction that the system of radiation and absorption is a *closed* system. The result is that Eq. (16–49) is required by the conservation of momentum. Conversely, the assumptions of conservation of momentum and the absence of an ether can be used to derive the mass-energy equivalence. In this way, the relation $E = mc^2$ can be obtained without introducing the entire relativistic kinematics.

16–5 The Minkowski force. The force $\mathbf{F} = d\mathbf{p}/dt$ in a given frame is not the spatial part of a four-vector. On the other hand, a quantity known as the *Minkowski force,*

$$F^i = \frac{dp^i}{ds} = \frac{d}{ds}(c\mathbf{p}, mc^2)$$
$$= \left[\frac{1}{\sqrt{1-u^2/c^2}}\frac{d\mathbf{p}}{dt}, \frac{1}{\sqrt{1-u^2/c^2}}\frac{d(mc)}{dt}\right], \tag{16–50}$$

is a contravariant four-vector. The components of F^i can be written as

$$F^i = \left(\frac{\mathbf{F}}{\sqrt{1-u^2/c^2}}, \frac{\mathbf{F}\cdot\frac{\mathbf{u}}{c}}{\sqrt{1-u^2/c^2}}\right), \tag{16–51}$$

where $\mathbf{F}$ and $\mathbf{u}$ are the ordinary three-dimensional force and velocity of Eqs. (16–39) and (16–40). (Note that the dimensions of F^i *are* those of a force!) The transformation laws for force can be derived from the four-vector character of F^i. We shall restrict ourselves to the case where $\mathbf{F}$ is proper in what we have called the Σ frame, i.e., $\mathbf{u} = 0$, $F^i = (\mathbf{F}, 0)$. In the Σ' frame $u'_x = v$, since $u_x = 0$, and we obtain

$$F'_x = F_x, \tag{16–52}$$

$$F'_y = F_y\sqrt{1-\beta^2}, \tag{16–53}$$

$$F'_z = F_z\sqrt{1-\beta^2}. \tag{16–54}$$

An interesting application of the transformation equations for $\mathbf{F}$ is afforded by the equilibrium of the right-angled lever shown in Fig. 16–3. To an observer at rest in Σ, the lever is in static equilibrium under the action of F_x and F_y as shown, so that

$$F_xL_y = F_yL_x. \tag{16–55}$$

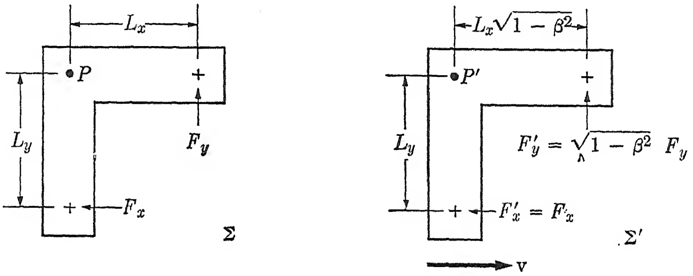


FIG. 16–3 Right-angled lever in a moving coordinate system.

To an observer at rest in Σ', we should expect the lever to remain in rotational and translational equilibrium, for otherwise the two inertial frames would be distinguishable. Using the transformation equations for lengths and forces, we find that to an observer in Σ' a net torque, T', is acting on the lever in a counterclockwise direction, where

$$T' = F_xL_y - (\sqrt{1-\beta^2}\,F_y)(\sqrt{1-\beta^2}\,L_x) = \frac{F_xL_yv^2}{c^2}. \qquad (16\text{–}56)$$

This torque results in no rotation, however, for F'_x is doing work on the lever at the rate $F'_xv = F_xv$, and the angular momentum of the lever is increasing at the rate

$$\frac{dM}{dt} = \frac{(F_xv)vL_y}{c^2} = \frac{F_xL_yv^2}{c^2}. \qquad (16\text{–}57)$$

Hence to an observer in Σ', even though neither the torque nor the change in angular momentum is zero, the existing torque exactly balances the gain in angular momentum and the equilibrium condition is preserved as an invariant property. (In this discussion we have omitted all mention of the mechanism by which forces are transmitted through the lever. The laws of elasticity are also profoundly modified by relativity: the lever cannot be treated as a rigid body, since the velocity of propagation of an impulse is limited. A more detailed discussion, taking account of these modifications, does not alter the above conclusions, however.)

Before leaving the subject of forces, let us consider briefly the motion of a particle under the influence of external forces. From

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = \frac{d}{dt}\left(\frac{m_0\mathbf{u}}{\sqrt{1-u^2/c^2}}\right), \qquad (16\text{–}39)$$

we obtain

$$\mathbf{F} = \frac{m_0}{\sqrt{1-u^2/c^2}}\frac{d\mathbf{u}}{dt} + \frac{m_0u\mathbf{u}}{c^2(1-u^2/c^2)^{3/2}}\frac{du}{dt}. \qquad (16\text{–}58)$$

This shows that the acceleration of a particle requires a component of force parallel to the velocity as well as that parallel to the acceleration. For low velocities the second member of the right side of Eq. (16–58) is much smaller than the first term. When u/c is very small in comparison with unity, Eq. (16–58) can be approximated by $\mathbf{F} = m_0\,d\mathbf{u}/dt$, the classical form of Newton's second law of motion.

16–6 The collision of two similar particles. Without reference to tensor methods the dependence of mass on velocity can be deduced by use of the Lorentz transformation and the requirement that in a two-particle collision momentum be conserved. This is an example of the *Gedanken Experimente* process of extending the principles of relativity.

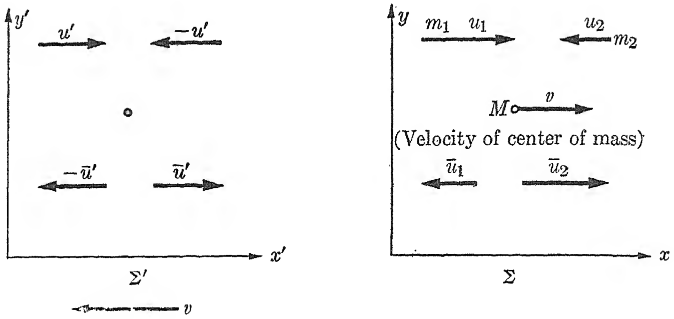


FIG. 16–4 Collision of two similar particles. The bars designate quantities after the collision.

Consider a head-on collision between two particles of equal properties. We may choose Σ' as the center of mass frame in which the collision is symmetrical, i.e., they appear to approach with equal velocities, as in Fig. 16–4. After a short period of coalescence they will move apart in opposite directions with velocities of equal magnitude in this frame.

We now postulate that momentum and mass are to be conserved in any frame. Hence, in Σ,

$$m_1 + m_2 = M, \tag{16–59}$$

$$m_1 u_1 + m_2 u_2 = Mv, \tag{16–60}$$

where M is the combined mass during collision. From the longitudinal velocity addition relations, we have

$$u_1 = \frac{u' + v}{1 + \dfrac{u'v}{c^2}}, \quad u_2 = \frac{-u' + v}{1 - \dfrac{u'v}{c^2}}. \tag{16–61}$$

By eliminating M from Eqs. (16–59) and (16–60), and making use of Eq. (16–61), we obtain

$$\frac{m_1}{m_2} = \frac{1 + u'v/c^2}{1 - u'v/c^2} = \frac{\sqrt{1 - u_2^2/c^2}}{\sqrt{1 - u_1^2/c^2}}. \tag{16–62}$$

Hence, in order to preserve the postulated conservation laws in all frames, we must have

$$m = \frac{m_0}{\sqrt{1 - u^2/c^2}}, \tag{16–63}$$

where m_0 is the mass in a frame proper to the particle. This is in agreement with Eq. (16–35), and the further deductions follow as above.

This description is of particular interest, since it enables us to obtain from Eq. (16–61) the mass M during the collision:

$$\begin{aligned} M &= m_0\left(\frac{1}{\sqrt{1-u_1^2/c^2}}+\frac{1}{\sqrt{1-u_2^2/c^2}}\right) \\ &= \frac{2m_0}{\sqrt{1-v^2/c^2}\sqrt{1-u'^2/c^2}}. \end{aligned} \tag{16–64}$$

This is larger than $2m_0/\sqrt{1-v^2/c^2}$, which would be the mass of the two particles of rest mass m_0 moving with velocity v. This increased mass represents the increase in energy of the two particles during collision owing to the stored elastic energy or to the energy increase where it is not all released again in kinetic form. The distinction between inelastic and elastic collisions therefore essentially disappears so far as the first part of a collision is concerned. If the two particles separate after the collision, a knowledge of the energy changes during the impact is necessary for the determination of the final velocities in any frame.

SUGGESTED REFERENCES

P. G. Bergmann, *An Introduction to the Theory of Relativity*, Chapters 5 and 6. The tensor calculus is presented in such a way that it will be applicable to the general theory of relativity.

R. C. Tolman, *Relativity, Thermodynamics, and Cosmology*, Chapters III and IV. Tolman and his collaborators have been responsible for such "thought experiments" as the collision of two similar particles.

L. Landau and E. Lifshitz, *The Classical Theory of Fields*, Chapters 1 and 2. We shall have further occasion to refer to this excellent book on relativistic electrodynamics. The authors stress rather formally the internal logic of the subject, and begin with the principle of relativity and the variation principle.

C. Møller, *The Theory of Relativity*, Chapter IV. Chapter VI includes the mechanics of elastic continua, but its study is better postponed until after the relativistic formulation of electrodynamics, Chapter V.

H. Goldstein, *Classical Mechanics*, Chapter 6. Many textbooks on mechanics include a section on relativity, and this is certainly one of the best.

Of the many excellent mathematical books on tensor calculus we recommend:

C. E. Weatherburn, *An Introduction to Riemannian Geometry and the Tensor Calculus*.

J. L. Synge and A. Schild, *Tensor Calculus*.

EXERCISES

1. Let a particle of initial energy W and rest energy W_0 hit a like particle at rest. Show that if $W \gg W_0$, the maximum energy available in the zero momentum frame is $\sqrt{2WW_0}$.

AGRICULTURAL

2. A π^0 meson of rest energy W_0 moving with velocity u in the laboratory, disintegrates into two γ-rays. Calculate (a) the energy distribution of γ-rays from π^0 mesons in the laboratory, (b) the distribution in angle of one of the γ-rays against the other in the laboratory.

3. If a particle of initial kinetic energy T_0 and rest energy W_0 strikes a like particle at rest, show that the kinetic energy T of the particle scattered at an angle θ is

$$T = \frac{T_0 \cos^2 \theta}{1 + \dfrac{T_0 \sin^2 \theta}{2W_0}}.$$

4. To produce a "proton pair" an energy of $2M_0c^2 = 1862$ Mev is required in the center-of-mass system. What is the minimum energy (a) of a proton, and (b) of an electron, needed to produce such an event by striking a proton at rest?

5. In experiments on proton-neutron scattering the scattered and recoil particles are at 90° to each other for nonrelativistic velocities. Calculate the first order correction (assumed small) to this angle for velocities large enough to make such a correction necessary.

CHAPTER 17

COVARIANT FORMULATION OF VACUUM ELECTRODYNAMICS

We recall that the Lorentz transformation was introduced by consideration of the propagation of an electromagnetic wave. Actually, the homogeneous equation governing electromagnetic wave propagation is already in covariant form, since the D'Alembertian operator

$$\Box = -\frac{\partial}{\partial x_i}\frac{\partial}{\partial x^i} \tag{17–1}$$

is an invariant. In general, Maxwell's equations and their consequences lend themselves very simply to covariant description. This follows from the fact that no modifications at all are necessary in the laws of electrodynamics to make them agree with the requirements of relativity. The covariant formulation of the space-time coordinates in the equations automatically puts the rest of the equation into covariant form.

17–1 The four-vector potential. Relativistically, it is clear that charge density and current are simply different aspects of the same thing. If we have a "proper" charge density ρ_0 in a frame where such charges are at rest, then the contravariant vector

$$j^i = \rho_0 \frac{dx^i}{ds} \tag{17–2}$$

has the components

$$j^i = \left(\rho\frac{\mathbf{u}}{c}, \rho\right), \tag{17–3}$$

where

$$\rho = \frac{\rho_0}{\sqrt{1 - u^2/c^2}}. \tag{17–4}$$

Hence the transformation equations for charge and current densities follow automatically. And since charge and current densities are components of a single four-vector j^i, we are led to combine the inhomogeneous wave equations, Eqs. (13–7) and (13–8), in an analogous way. Expressed as

$$\Box\phi = -\rho/\epsilon_0, \tag{17–5}$$

$$\Box c\mathbf{A} = -\rho\mathbf{u}/c\epsilon_0, \tag{17–6}$$

these equations are equivalent to

$$\Box\phi^i = -\frac{j^i}{\epsilon_0}, \tag{17-7}$$

where

$$\phi^i = (c\mathrm{A},\phi). \tag{17-8}$$

The scalar and vector potentials are therefore no longer quantities permitting independent description; they are different aspects of the same thing. It follows that the same can be said of electric and magnetic fields: in some sense they, too, are different aspects of the same thing.

The equation of continuity of charge and current takes the simple covariant form

$$\frac{\partial j^i}{\partial x^i} = 0, \tag{17-9}$$

and the Lorentz condition, Eq. (13–3), becomes its counterpart:

$$\frac{\partial \phi^i}{\partial x^i} = 0. \tag{17-10}$$

In terms of the covariant components of ϕ^i, the gauge transformations, Eqs. (13–9) and (13–10), combine as the single equation

$$\phi_i' = \phi_i + \frac{\partial}{\partial x^i}(c\psi), \tag{17-11}$$

giving both the correct transformation character and the correct sign. The derivation of the field from the ϕ^i and laws of physical consequence must not depend on the choice of the scalar function ψ.

Equation (17–4), which gives the transformation from a charge density at rest to a charge density in a nonproper frame, is such as to ensure the invariance of total charge. A spatial volume element dv is related to a proper spatial volume dv_0 by

$$dv = dv_0\sqrt{1 - u^2/c^2}, \tag{17-12}$$

since only one dimension suffers a Lorentz contraction. Hence,

$$\rho\, dv = \rho_0\, dv_0 \tag{17-13}$$

and the charge within a given boundary remains invariant. The electronic charge e thus remains a universal constant in the theory of relativity. Since no charges have been found in nature which are not integral multiples of e, total charge could be measured by a counting operation which is presumably also an invariant. These facts are all in agreement with experiment. The invariance of total charge is also a direct consequence of the theorem stated in Eq. (16–22) combined with Eq. (17–9).

Let us obtain the integral of the inhomogeneous wave equation, Eq. (17–7), corresponding to an (invariant) point charge e at a point $(x^j)_1$. We know that in a proper frame, as in Fig. 17–1, the integral of Eq. (17–7) is simply the Coulomb potential,

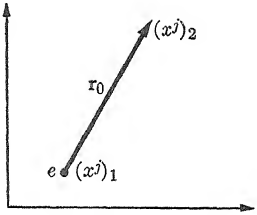


FIG. 17–1 Source point and field point in a proper frame.

$$\phi_0^i = \left(0, \frac{e}{4\pi\epsilon_0}\frac{1}{r_0}\right), \tag{17–14}$$

where $\mathbf{r}_0$ is the proper vector distance from the source to the field point. The potential signal at $(x^j)_2$ is to be measured at the time corresponding to the retardation condition

$$R^jR_j = 0 = -r^2 + c^2t^2, \tag{17–15}$$

where

$$R^j = (x^j)_1 - (x^j)_2 = (\mathbf{r},ct) = (\mathbf{r},r). \tag{17–16}$$

To make Eq. (17–14) valid in any frame, we seek to write the potential solution in tensor form such that it reduces to the Coulomb potential if $\mathbf{u} = 0$. In a proper frame, the four-velocity, $u^i = dx^i/ds$, has the components

$$u_0^i = (0,1). \tag{17–17}$$

Now the invariant, u^iR_i, can be evaluated in the proper frame where $R_0^i = (\mathbf{r}_0,r_0)$, and is just

$$u^iR_i = r_0. \tag{17–18}$$

Equation (17–14) can then be written in tensor form,

$$\phi^i = \frac{e}{4\pi\epsilon_0}\frac{u^i}{u^jR_j}, \tag{17–19}$$

subject to the condition $R_jR^j = 0$. This equation is now valid in any frame, whether proper or not.

In order to write Eq. (17–19) in terms of three-dimensional vectors, we recall that

$$u^j = \left(\frac{\mathbf{u}}{c\sqrt{1 - u^2/c^2}}, \frac{1}{\sqrt{1 - u^2/c^2}}\right). \tag{16–29}$$

Hence,

$$u^jR_j = \frac{-\mathbf{r}\cdot\mathbf{u}}{c\sqrt{1 - u^2/c^2}} + \frac{r}{\sqrt{1 - u^2/c^2}}, \tag{17–20}$$

and Eq. (17–19) then has the components

$$\phi^i = \frac{e}{4\pi\epsilon_0}\left(\frac{\mathbf{u}}{cs}, \frac{1}{s}\right), \tag{17–21}$$

where

$$s = r - \frac{\mathbf{r}\cdot\mathbf{u}}{c}. \tag{17–22}$$

The potentials of Eq. (17–21) are called the Liénard-Wiechert potentials, and later we shall derive them from the retarded potentials of Chapter 13. In Eq. (17–21) **u** enters as the velocity of the observer relative to the frame in which the charge was at rest at the time of "emission" of the signal.

17–2 The electromagnetic field tensor. In writing down the covariant derivation of the fields from the potentials a problem arises which has its origin in ordinary three-dimensional vector analysis: **E** and **B** are not vectors of the same kind. The prototype for an ordinary vector is the displacement of a point in space. Such vectors are called *polar* vectors, and **u**, **F**, **E**, **D**, etc., are of this type. The scalar products of these vectors are unchanged by all orthogonal transformations of the coordinates—translations, rotations, and reflections—although if any or all of the coordinate axes are changed in sign the corresponding vector components change sign. The same cannot be said of the "vector product" of two polar vectors: the very definition of vector product depends on whether **A** × **B** is to be taken in the right-handed or left-handed sense. (The usual prototype for a vector product represents the area of the parallelogram defined by **A** and **B**, but the "sense" of such a surface is arbitrary.) It is clear that if all three coordinate components of **A** and **B** are changed in sign no change at all occurs in **A** × **B**. Quantities that transform like vector products are called *axial* vectors, and examples are torque, angular momentum, **B**, **H**, and **M**. They transform like ordinary vectors under translations and proper rotations of the coordinate system, but under any changes involving reflections of the coordinates the relative sign between axial and polar vectors is reversed. The scalar product of an axial vector and a polar vector is not a true scalar, but a pseudoscalar: it changes sign under an inversion of the axes. Physical vector relations can, of course, equate only vectors of the same kind.

Axial vectors cannot form the spatial components of a four-vector. On the other hand, the components of **C** = **A** × **B** can be expressed by two indices, those of the components of **A** and **B**:

$$C_{\alpha\beta} = A_\alpha B_\beta - A_\beta B_\alpha = -C_{\beta\alpha}, \tag{17–23}$$

where

$$C_{12} = C_z, \quad C_{23} = C_x, \quad C_{31} = C_y, \tag{17–24}$$

in a right-handed Cartesian coordinate system. For example, the equation $\nabla \times \mathbf{E} = -\dot{\mathbf{B}}$ is written in this notation as

$$\frac{\partial E_\beta}{\partial x^\alpha} - \frac{\partial E_\alpha}{\partial x^\beta} = -\dot{B}_{\alpha\beta}, \tag{17–25}$$

and $\nabla \times \mathbf{A} = \mathbf{B}$ is written as

$$\frac{\partial A_\beta}{\partial x^\alpha} - \frac{\partial A_\alpha}{\partial x^\beta} = B_{\alpha\beta}, \tag{17–26}$$

where α and β are restricted to 1, 2, 3, and ordinary three-dimensional components are involved on the left side of the equations. On the other hand, $\mathbf{E}$ is polar; it can be derived from the potentials by

$$E_\alpha = -\frac{\partial \phi}{\partial x^\alpha} - \frac{\partial A_\alpha}{\partial t}, \qquad (\alpha = 1, 2, 3). \tag{17–27}$$

Examination of Eqs. (17–26) and (17–27) in the light of tensor requirements leads us to introduce a four-dimensional antisymmetric field tensor F_{ij} which, as a function of the covariant four-vector potential

$$\phi_i = (-c\mathbf{A}, \phi), \tag{17–28}$$

is

$$F_{ij} = \frac{\partial \phi_j}{\partial x^i} - \frac{\partial \phi_i}{\partial x^j}. \tag{17–29}$$

The contravariant tensor would be given by

$$F^{ij} = \frac{\partial \phi^j}{\partial x_i} - \frac{\partial \phi^i}{\partial x_j}. \tag{17–30}$$

In conformity with Eqs. (17–26) and (17–27), the components of F_{ij} are

$$F_{ij} = \begin{array}{cc} & j \rightarrow \\ \begin{array}{c} i \\ \downarrow \end{array} & \begin{pmatrix} 0 & -cB_z & +cB_y & -E_x \\ +cB_z & 0 & -cB_x & -E_y \\ -cB_y & +cB_x & 0 & -E_z \\ +E_x & +E_y & +E_z & 0 \end{pmatrix} \end{array}. \tag{17–31}$$

The components of the corresponding contravariant tensor F^{ij} are

$$F^{ij} = \begin{array}{cc} & j \rightarrow \\ \begin{array}{c} i \\ \downarrow \end{array} & \begin{pmatrix} 0 & -cB_z & +cB_y & +E_x \\ +cB_z & 0 & -cB_x & +E_y \\ -cB_y & +cB_x & 0 & +E_z \\ -E_x & -E_y & -E_z & 0 \end{pmatrix} \end{array} = g^{in} g^{jm} F_{nm}. \tag{17–32}$$

Since it is possible to write the fields covariantly in terms of the potential, it should be possible to write Maxwell's equations themselves in covariant form. It is easily verified that the source equations,

$$\begin{gathered} \nabla \cdot \mathbf{E} = \rho/\epsilon_0, \\ \nabla \times \mathbf{B} = \mu_0 \left(\rho \mathbf{u} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right), \end{gathered} \tag{17-33}$$

correspond to

$$\frac{\partial F^{ji}}{\partial x^j} = \frac{j^i}{\epsilon_0}, \tag{17-34}$$

which is just the covariant divergence of the field tensor. Agreement of Eq. (17-34) with the equation of continuity, Eq. (17-9), is obvious from the antisymmetric character of the field tensor. The other field equations are a little more awkward:

$$\nabla \cdot \mathbf{B} = 0 \quad \text{and} \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{17-35}$$

correspond to

$$\frac{\partial F_{ij}}{\partial x^k} + \frac{\partial F_{jk}}{\partial x^i} + \frac{\partial F_{ki}}{\partial x^j} = 0. \tag{17-36}$$

The left side of Eq. (17-36) will vanish identically unless $i \neq j \neq k$. Moreover, permutation of indices will not change the content, and hence only four of the sixty-four equations formally represented by Eq. (17-36) are nontrivial.

Equation (17-36) can be written in a form resembling Eq. (17-34) (which is equivalent to a four-vector) by taking account of the fact that it is a completely antisymmetric tensor of third rank. The completely antisymmetric tensor of fourth rank, P^{ijkl}, is defined so that its components are zero unless $i \neq j \neq k \neq l$, and equal to ± 1 according to whether $ijkl$ is an even or odd permutation of 1234. P^{ijkl} transforms like a tensor under translations and proper rotations of the coordinates, but not under reflections or inversions (improper rotations): changing the sign of an odd number of coordinate axes does not change the components of P^{ijkl} under this definition, contrary to what is expected of a tensor. It is thus not strictly a tensor, but a *pseudotensor*. The pseudotensor formed by multiplying the antisymmetric field tensor by P^{ijkl} is called its *dual*:

$$G^{ij} = P^{ijkl} F_{kl}. \tag{17-37}$$

It can be seen that G^{ij} is constructed like F^{ij} except that **E** and $c\mathbf{B}$ are interchanged. In terms of G^{ij}, Eq. (17-36) is equivalent to

$$\frac{\partial G^{ij}}{\partial x^i} = 0. \tag{17-38}$$

This equation is written so as to permit the introduction of magnetic poles and pole currents if they exist. Since G^{ij} is a pseudotensor, any hypothetical pole four-vector to be added to Eq. (17–38) is a pseudovector.

The tensor expression for the fields immediately permits a derivation of the Lorentz transformation of the fields. From Eq. (16–12),

$$F^{ij\prime} = Q^i_k Q^i_l F^{kl}, \tag{17–39}$$

and thence we can easily derive the relations

$$E'_{\parallel} = E_{\parallel}, \tag{17–40}$$

$$B'_{\parallel} = B_{\parallel}, \tag{17–41}$$

$$\mathbf{E}'_{\perp} = \gamma(\mathbf{E}_{\perp} + \mathbf{v} \times \mathbf{B}_{\perp}), \tag{17–42}$$

$$\mathbf{B}'_{\perp} = \gamma(\mathbf{B}_{\perp} - \mathbf{v} \times \mathbf{E}_{\perp}/c^2), \tag{17–43}$$

where $\gamma = 1/\sqrt{1 - \beta^2}$, as in Chapter 16, and $E_{\parallel}$, $B_{\parallel}$ and $\mathbf{E}_{\perp}$, $\mathbf{B}_{\perp}$ are the components of $\mathbf{E}$ and $\mathbf{B}$ parallel and normal to $\mathbf{v}$, respectively.

Equations (17–42) and (17–43) can be interpreted physically in a fairly simple way. The terms linear in $\mathbf{v}$ (except for the factor γ) should be essentially classical, i.e., describable by Maxwell's equations without explicit use of relativistic arguments. Equation (17–42) corresponds to the fact that to order β^2 a particle moving relative to a magnetic field experiences an electric field $\mathbf{E} + \mathbf{v} \times \mathbf{B}$, as has been discussed in detail in Chapter 9. To interpret Eq. (17–43), we may consider the flux through a rectangle moving through a region containing an electric field, as shown in Fig. 17–2. The rectangle extends outside the region where $\mathbf{E}$ is different from zero. In exact analogy to the magnetic case discussed in Section 9–3, the line integral of the magnetic field $\mathbf{B}'$ seen by an observer moving with the rectangle would be governed by the rate of change of electric flux through this rectangle as well as the real current J linking the rectangle. The rate of change of electric flux is given by

$$\oint \mathbf{E} \cdot (\mathbf{v} \times d\mathbf{l} = -\oint (\mathbf{v} \times \mathbf{E}) \cdot d\mathbf{l}. \tag{17–44}$$

Therefore the circulation of $\mathbf{B}'$ is given by

$$\oint (\mathbf{B} \cdot d\mathbf{l}) = \frac{1}{c^2}\frac{d}{dt}\int \mathbf{E} \cdot d\mathbf{S} + \mu_0 J = -\frac{1}{c^2}\oint (\mathbf{v} \times \mathbf{E}) \cdot d\mathbf{l} + \mu_0 J. \tag{17–45}$$

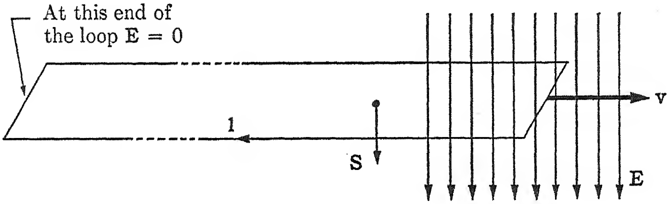


FIG. 17–2 Rectangular path of integration moving through an inhomogeneous electric field.

In terms of the rest coordinates, this is equivalent to

$$\mathbf{B} = \mathbf{B}' + \frac{1}{c^2}(\mathbf{v} \times \mathbf{E}), \tag{17–46}$$

in agreement with Eq. (17–43) for small β.

The transformation equations for the fields are of considerable value in the solution of practical problems involving the motion of electrons and ions in electromagnetic fields. It is frequently possible to transform away either the electric or the magnetic field by choosing a suitable Lorentz frame, whereupon the solution may be much simplified.

17–3 The Lorentz force. The Lorentz force per unit volume, $\mathbf{f} = \rho(\mathbf{E} + \mathbf{u} \times \mathbf{B})$, is the space part of a four-vector,

$$f^i = F^{ik} j_k = \left(\mathbf{f}, \frac{\mathbf{u} \cdot \mathbf{f}}{c}\right). \tag{17–47}$$

The fourth component of f^i is $1/c$ times the power expended by the electric field per unit volume. But the total force acting on a charge occupying a volume δv, so that $\delta q = \rho\ \delta v$, is not the space part of a four-vector, since it is given by

$$\mathbf{F} = \mathbf{f}\ \delta v = \delta q\ (\mathbf{E} + \mathbf{u} \times \mathbf{B}). \tag{17–48}$$

In a proper frame, since δq is an invariant,

$$\mathbf{F}_0 = \delta q\ \mathbf{E}_0 \tag{17–49}$$

and hence, in general, from Eqs. (17–40) and (17–42),

$$F_{||} = F_{0||}, \tag{17–50}$$

$$\mathbf{F}_{\perp} = \gamma^{-1}\mathbf{F}_{0\perp}. \tag{17–51}$$

These equations are in agreement with the mechanical force transformations, Eqs. (16–52), (16–53), (16–54). It follows that equilibrium between mechanical forces and electrical forces is invariant to the choice of frame—the nature of the force does not affect its transformation properties. Moreover, we are justified in defining force by the relation $\mathbf{F} = d\mathbf{p}/dt$ if for $\mathbf{F}$ we use the Lorentz force, Eq. (17–48), and thus Eq. (16–50) is the relativistically correct expression for the equation of motion of a charged particle in an electromagnetic field. In Chapter 23, we shall return to consider the covariant Hamiltonian formulation of the equations both for particles and for fields.

SUGGESTED REFERENCES

Excellent accounts will be found in the previously listed books by Becker, Bergmann, Møller, and Tolman. Landau and Lifshitz approach the subject from the behavior of a particle in a field. In addition we should list the following works.

A. Sommerfeld, *Electrodynamics*, Part III, Theory of Relativity and Electron Theory. Electrodynamics is the third volume of Professor Sommerfeld's justly famous Lectures on Theoretical Physics.

A. Einstein, a translation of the original paper On the Electrodynamics of Moving Bodies in *The Principle of Relativity.*

W. Pauli, *Relativitätstheorie*, Encyclopädie der Mathematischen Naturwissenschaften, Vol. V, or in special reprint. A comprehensive classical reference on the subject since 1921.

EXERCISES

Show that:

1. If **E** and **B** are perpendicular in one Lorentz frame, they are perpendicular in all Lorentz frames.
2. If $|\mathbf{E}| > |c\mathbf{B}|$ in any one Lorentz frame, then $|\mathbf{E}| > |c\mathbf{B}|$ in any other Lorentz frame, and vice versa.
3. If the angle between **E** and **B** is acute (or obtuse) in one Lorentz frame, it is acute (or obtuse) in any other Lorentz frame.
4. If **E** is perpendicular to **B** but $|\mathbf{E}| \neq |c\mathbf{B}|$, then there is a frame in which the field is either purely electric or purely magnetic.

CHAPTER 18

THE LIÉNARD-WIECHERT POTENTIALS AND THE FIELD OF A UNIFORMLY MOVING ELECTRON

The mathematical "machinery" of covariance is a powerful tool for deriving the consequences of electromagnetic theory, as we have seen, but the resulting expressions must be translated into the ordinary space and time variables of the observer in order to be compared with experiment. Most of the results of classical radiation theory were derived before the advent of relativity theory, it being assumed that the frame of the observer was one in which Maxwell's equations are valid. It is instructive, in view of the physical interpretation of the formulas, to go through some of these derivations, especially now that we have seen from relativistic considerations that their validity is much more general than could have been originally supposed.

18–1 The Liénard-Wiechert potentials. Let us consider the application of the retarded potentials, Eqs. (13–24) and (13–25), to compute the radiation from an electron. Now in classical electrodynamics the only thing known about the electron is that it has a certain total charge, and any calculation of its radiation field cannot involve details of how this charge may be distributed geometrically within the electron. On the other hand, it is impossible to assume that the charge has zero physical extent without introducing various mathematical divergences. But certain features of the radiation field are actually independent of the radius of the electron, provided only that it is small compared with the other dimensions of the radiation field. In our discussion of the electron and its behavior we shall assume that it has a finite radius, but we shall ascribe physical significance only to those properties which are independent of the magnitude of the radius. Actually a definite electron radius has no meaning as an observable quantity. It can be shown quantum mechanically that no length measurement on the electron can yield its radius. Moreover, as we have noted previously, a "rigid" dimension for an object is relativistically self-contradictory.

Prerelativistically, considerable care had to be exercised in applying the concept of retarded potential to a system whose total charge is known. If $[\rho]$ is the retarded charge density to be substituted in Eq. (13–25) to give the correct potential $\phi(x_\alpha,t)$, then $\int[\rho]\,dv$ does *not* in general represent the total charge of the system. The reason is that the various contributions to the integrand $[\rho]\,dv$ are evaluated at different times, and during

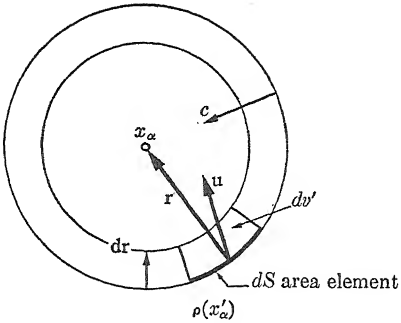


FIG. 18-1 Element dS of sphere collecting information from charge in motion with velocity **u**.

the time the information-collecting sphere of Fig. 18-1 sweeps over the charge distribution the charges may move so as to appear more or less dense than they should to give a correct value for the total charge. This can be illustrated by a detailed consideration of the process. Consider the sphere of Fig. 18-1 converging onto the point of observation x_α with a velocity c, and let it gather information as to the charge density within a certain charge system as it sweeps across the system. If the charge system has an average velocity component in the same direction as the motion of the converging sphere the volume integral of the retarded charge density will be in excess of the total charge. If the charge distribution has an average velocity component in opposition to the velocity of the contracting sphere the integral will give a result less than the total charge of the system.

The situation is analogous to the problem of taking a census of the population of a country. Let us assume that a group of census takers converges upon the "information center" with a certain speed, measuring the population density by counting the people every day as they travel. The correct population will differ from the total of the census takers' information depending on whether the population had a net migration trend with the census takers (in which case the true population is less than the sum of the reported densities) or against the census takers. Similarly, the retarded potential of an approaching charge will be larger than that of a receding charge at the same distance from the observer, since the approaching charge stays longer within the information-collecting sphere.

Let us consider the radiation field of an electron whose velocity is comparable to c. We shall assume that our electron is a system about which we know (a) that the total charge is e, and (b) that within an unspecified but small volume all parts of the electron's charge are moving systematically with a velocity **u**. Let **r** be the radius vector from a point where there is charge to the point of observation on which the sphere of Fig. 18-1 is converging; **r** is the "retarded" value of the radius, measured at the time of passage of the sphere. If the electron is at rest, the amount of charge that the sphere will cross during the time dt as it shrinks by dr is given by $[\rho]\, dS\, dr$. On the other hand, if **u** is different from zero a quantity of charge which is less than $[\rho]\, dS\, dr$ by the amount $[\rho]\, dS(\mathbf{u} \cdot \mathbf{r}/r)\, dt$ will be crossed by the sphere. Thus the amount of charge observed by this means

AGRICULTURAL

of collecting information is

$$de = [\rho]\,dv' - [\rho]\frac{\mathbf{u}\cdot\mathbf{r}}{r}\,dS\,dt. \tag{18–1}$$

But $dt = dr/c$, and $dS\,dr = dv'$, so that

$$de = [\rho]\,dv' - [\rho]\frac{\mathbf{u}\cdot\mathbf{r}}{rc}\,dv'. \tag{18–2}$$

Solving for the retarded charge density, we obtain

$$[\rho]\,dv' = \frac{de}{1 - \dfrac{\mathbf{u}\cdot\mathbf{r}}{cr}}$$

or

$$\frac{[\rho]\,dv'}{r} = \frac{de}{r - \dfrac{\mathbf{u}\cdot\mathbf{r}}{c}}. \tag{18–3}$$

The retarded potentials of Eqs. (13–25) and (13–24) thus become

$$\phi = \frac{1}{4\pi\epsilon_0}\int\frac{de}{r - \dfrac{\mathbf{r}\cdot\mathbf{u}}{c}}, \tag{18–4}$$

$$\mathbf{A} = \frac{\mu_0}{4\pi}\int\frac{\mathbf{u}\,de}{r - \dfrac{\mathbf{r}\cdot\mathbf{u}}{c}}. \tag{18–5}$$

At the limit of a point charge the distance-dependent terms are slowly varying, and since $\int de = e$, the known electronic charge, the potentials of a point charge are

$$\phi = \frac{1}{4\pi\epsilon_0}\left[\frac{e}{r - \dfrac{\mathbf{r}\cdot\mathbf{u}}{c}}\right], \tag{18–6}$$

$$\mathbf{A} = \frac{\mu_0}{4\pi}\left[\frac{e\mathbf{u}}{r - \dfrac{\mathbf{r}\cdot\mathbf{u}}{c}}\right]. \tag{18–7}$$

Equations (18–6) and (18–7) are known as the Liénard-Wiechert potentials of a single electron. Note that these expressions are independent of the extent of the charge, and therefore independent of any detailed electronic model. They are, in fact, just Eqs. (17–21), but the significance of the velocity is quite different. Here the velocity is that of the charge in its retarded position relative to the special coordinate frame in which it had to be assumed, prior to relativity theory, that Maxwell's equations were valid. In Chapter 17 we saw that the velocity was that of the observer relative to the frame in which the charge was at rest at the time the signal was emitted. All the detailed calculations of fields from moving charges made on the assumption that there is one frame in which the wave equation is correct are equivalent to those resulting from a covariant formulation of electrodynamics if we thus re-interpret the velocity.

18–2 The fields of a charge in uniform motion. Since the relation of the "retarded" position to the "present" position of an electron is not, in general, known, the Liénard-Wiechert potentials ordinarily permit only the evaluation of the fields in terms of the retarded positions and velocities of the charges. If the motion is uniform, however, it is possible to express the potentials and the fields in terms of the "present" position of the charge, i.e., the position of the charge at time t that the information-collecting sphere of Fig. 18–1 converges at the point of observation.

Consider an electron, as in Fig. 18–2, moving with a uniform velocity in the x-direction. Let us evaluate the Liénard-Wiechert denominator $s = r - (\mathbf{r} \cdot \mathbf{u})/c$ in terms of the present position of the electron. Since by the geometry of the figure, $\mathbf{r}_0 \times \mathbf{u} = \mathbf{r} \times \mathbf{u}$, it can be easily verified that

$$s^2 = r_0^2 - \left(\frac{\mathbf{r}_0 \times \mathbf{u}}{c}\right)^2. \tag{18–8}$$

The denominator s can then be expressed explicitly in terms of the present position coordinates x_0, y_0, z_0, or in terms of r_0 and the angle ψ between u and r_0:

$$\begin{aligned} s &= \sqrt{x_0^2 + y_0^2 + z_0^2 - \frac{u^2}{c^2}(y_0^2 + z_0^2)} \\ &= \sqrt{x_0^2 + \left(1 - \frac{u^2}{c^2}\right)(y_0^2 + z_0^2)} \\ &= r_0 \sqrt{1 - \frac{u^2}{c^2}\sin^2\psi}. \end{aligned} \tag{18–9}$$

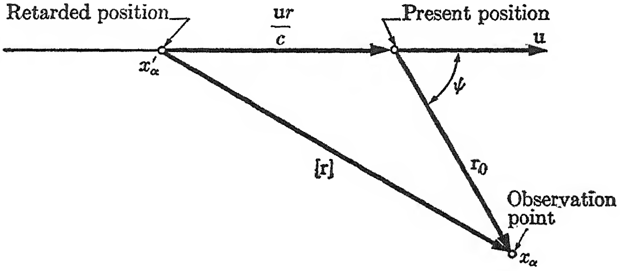


FIG. 18–2 Potential and field coordinates of an electron in uniform motion.

With this substitution for the denominator of Eqs. (18–6) and (18–7) the Liénard-Wiechert potentials are given in terms of the present coordinates of the electron.

The fields are easy to compute explicitly in this case. The electric field is

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}. \tag{13–1}$$

The time derivative can be evaluated in terms of the spatial derivative by noting that the field must be carried by the uniformly moving charge. A stationary observer will see the same change in the fields during the time dt as will an observer who moves a distance $-\mathbf{u}\,dt$ from the same position without any lapse of time. Hence the time derivative can be replaced by

$$\frac{\partial}{\partial t} = -u\frac{\partial}{\partial x}. \tag{18–10}$$

With the use of this relation in Eq. (13–1), the components of the electric field become simply:

$$E_x = \frac{ex_0}{4\pi\epsilon_0 s^3}\left(1 - \frac{u^2}{c^2}\right), \tag{18–11}$$

$$E_y = \frac{ey_0}{4\pi\epsilon_0 s^3}\left(1 - \frac{u^2}{c^2}\right), \tag{18–12}$$

$$E_z = \frac{ez_0}{4\pi\epsilon_0 s^3}\left(1 - \frac{u^2}{c^2}\right). \tag{18–13}$$

Although $\mathbf{A}$ has only an x-component, the electric field is symmetrical (except for the denominator) in its three components. Note that the field is directed toward the "present" position of the electron (for a negative electron), and not toward the "retarded" position. Vectorially the electric field is given by

$$\mathbf{E} = \frac{e}{4\pi\epsilon_0 s^3}\left[\mathbf{r} - \frac{\mathbf{u}r}{c}\right]_{\text{ret}}\left(1 - \frac{u^2}{c^2}\right) = \frac{e\mathbf{r}_0(1 - u^2/c^2)}{4\pi\epsilon_0 s^3}$$

$$= \frac{e\mathbf{r}_0}{4\pi\epsilon_0 r_0^3}\,\frac{\left(1 - \dfrac{u^2}{c^2}\right)}{\left(1 - \dfrac{u^2}{c^2}\sin^2\psi\right)^{3/2}}, \tag{18–14}$$

which shows explicitly the direction of the field.

The magnetic field of the uniformly moving electron, as computed from

$\mathbf{B} = \nabla \times \mathbf{A}$, turns out to be

$$\mathbf{B} = \frac{\mu_0 e}{4\pi} \frac{\mathbf{u} \times \mathbf{r}_0}{s^3} \left(1 - \frac{u^2}{c^2}\right) = \frac{1}{c^2} \mathbf{u} \times \mathbf{E}. \tag{18-15}$$

For low velocities, $u \to 0$, $s \to r$, Eqs. (18–14) and (18–15) reduce to the Coulomb and Biot-Savart fields. For high velocities, $u \to c$, both field magnitudes depend on the angle between the direction of motion of the electron and the radius vector $\mathbf{r}_0$. From Eq. (18–14) it is seen that the electric field is increased in a direction at right angles to the direction of motion in the ratio of $1/\sqrt{1 - u^2/c^2}$, while in the direction of motion the field is decreased in the ratio $(1 - u^2/c^2)$. At very high velocities the field thus resembles more and more the field in a plane wave. For a short time, as a high-velocity electron passes an observer, he sees a purely transverse electric and magnetic field. Note, however, that a uniformly moving charge is nonradiating in the sense that its field does not represent an energy loss. This can be shown by direct evaluation of the Poynting vector corresponding to the fields given by Eqs. (18–14) and (18–15).

These formulas are identical with those obtained by applying a Lorentz transformation to the fields of a static charge, with no restriction on the velocity other than that implied by the limit $\beta \to 1$. (The proof of this statement is left to a problem.) Prior to relativity theory, however, considerable care had to be exercised in the application of these results, and the assumption that the observer is at rest with respect to the particular frame of reference in which Maxwell's equations were assumed valid was very troublesome. The theory of relativity has eliminated these difficulties by justifying the interpretation of $\mathbf{u}$ as the relative velocity between the electron and an observer.

18–3 Direct solution of the wave equation. It is instructive to see how our conclusions regarding the field of a uniformly moving charge can be derived from the inhomogeneous wave equations. Equations (13–7) and (13–8) are to be solved subject to the subsidiary Lorentz condition of Eq. (13–3). The field of an electron moving with uniform velocity must be carried convectively along with the electron, which implies that the time and space derivatives are not independent. This fact can be expressed mathematically by a generalization of Eq. (18–10):

$$\frac{\partial}{\partial t} = -\mathbf{u} \cdot \nabla. \tag{18-16}$$

This means that any field parameter at a given point changes by the same amount in a time dt as at a fixed time it would differ from the same field parameter evaluated at a distance $-u\,dt$ along the direction of motion of the electron.

Let us consider a single component of the inhomogeneous equation

$$\Box\psi(x_\alpha,t) = -g(x_\alpha) \tag{13–11}$$

under the condition that the source charge moves with constant velocity in the x-direction. Then, by Eq. (18–16), the wave equation becomes

$$\left(1-\frac{u^2}{c^2}\right)\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}+\frac{\partial^2\psi}{\partial z^2} = -g(x,y,z). \tag{18–17}$$

This equation can be transformed by a change of variables

$$x_1 = \frac{x}{\sqrt{1-u^2/c^2}}; \quad y_1 = y; \quad z_1 = z; \tag{18–18}$$

to a simple electrostatic Poisson equation,

$$\nabla_1^2\psi = -g\left(\sqrt{1-u^2/c^2}\,x_1,y_1,z_1\right), \tag{18–19}$$

of which the solution is the ordinary Coulomb potential,

$$\psi(x_{\alpha_1}) = \frac{1}{4\pi}\int\frac{g(x'_{\alpha_1})\,dv'_1}{r(x'_{\alpha_1},x_{\alpha_1})}. \tag{18–20}$$

If we transform back to the original variables, this solution becomes

$$\psi(x_\alpha) = \frac{1}{4\pi}\int\frac{g(x'_\alpha)}{s}\,dv', \tag{18–21}$$

where

$$s = \sqrt{(x-x')^2+(1-u^2/c^2)[(y-y')^2+(z-z')^2]}. \tag{18–22}^*$$

Explicitly, then, for a charge e moving with a velocity $\mathbf{u}$,

$$\phi = \frac{e}{4\pi\epsilon_0 s}, \tag{18–23}$$

$$\mathbf{A} = \frac{e\mathbf{u}}{4\pi\epsilon_0 c^2 s}. \tag{18–24}$$

Equations (18–23) and (18–24) represent the same potentials as those obtained from the Liénard-Wiechert expressions. In this derivation, however, the question regarding the propagation velocity of the corresponding wave and the relation between present and retarded potentials does not enter, since by a suitable transformation we have reduced the equation to be solved, Eq. (13–11), to the static equation, Eq. (18–19). It is obvious that the purely mathematical process of Eq. (18–18) is in reality a Lorentz transformation, in which we transform the observer's position to a frame that is at rest relative to the electron whose field is to be computed.

---

* Note that $x - x'$ in Eq. (18–22) corresponds to the "present" coordinate x_0 in Eq. (18–9), and likewise for the other coordinates.

18–4 The "convection potential." The force which would be exerted by these fields on another electron moving with a velocity u parallel to that of the original electron producing the field is presumably given by the Lorentz expression:

$$\mathbf{F} = e(\mathbf{E} + \mathbf{u} \times \mathbf{B}).$$

If the fields are computed from the potentials of Eqs. (18–23) and (18–24), this force is given by

$$\mathbf{F} = \frac{e^2}{4\pi\epsilon_0}\left[-\nabla\left(\frac{1}{s}\right) + (\mathbf{u}\cdot\nabla)\frac{\mathbf{u}}{c^2 s} + \frac{\mathbf{u}}{c^2} \times \left(\nabla \times \frac{\mathbf{u}}{s}\right)\right], \tag{18–25}$$

which, by expansion of the vector product, becomes

$$\mathbf{F} = -\frac{e^2}{4\pi\epsilon_0}\nabla\left(\frac{1 - u^2/c^2}{s}\right). \tag{18–26}$$

Equation (18–26) can be written in the form

$$\mathbf{F} = -\nabla\psi, \tag{18–27}$$

where

$$\psi = \frac{e^2(1 - u^2/c^2)}{4\pi\epsilon_0 s} \tag{18–28}$$

is called the convection potential. The force of one electron on the other is thus derivable from a scalar potential, ψ, but this scalar potential does not have spherical symmetry about the position of the electron producing the field. In particular, since the direction of the force must be perpendicular to the surface of equal convection potential, we would conclude that the force $\mathbf{F}_2$ exerted by the electron e_1 at (x_1,y_1,z_1) on the electron e_2 at (x_2,y_2,z_2) is perpendicular to the ellipsoid

$$\begin{aligned} s &= \sqrt{(x_1 - x_2)^2 + (1 - u^2/c^2)[(y_1 - y_2)^2 + (z_1 - z_2)^2]} \\ &= \text{constant}, \end{aligned} \tag{18–29}$$

as shown in Fig. 18–3.

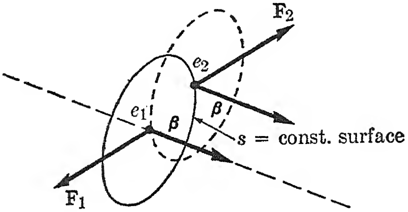


FIG. 18–3 Showing forces between two point charges as derived from the convection potential.

On the other hand, the reaction force $\mathbf{F}_1$ on the electron e_1 is perpendicular to the corresponding ellipsoid (shown by the dashed line in the figure) referred to the co-moving electron e_2. Hence, except when the line between the electrons is parallel or perpendicular to the direction of motion, the forces of action and reaction do not appear to be collinear. Therefore if the two electrons were connected by a rigid bar there would be a couple acting about an axis perpendicular to the plane of the line joining the electrons and the direction of motion. This will be recognized as the torque also predicted by Ampere's law when current elements are substituted for the moving charges, and which Trouton and Noble attempted to measure. The paradox produced by the observation of a null effect indicates the difficulties in interpreting the velocity of moving charges in prerelativistic electrodynamics.

The torque predicted here is real enough to an observer moving with a velocity $\mathbf{u}$ relative to the two charges, and should in that case be measurable if there were no mechanical considerations involved. We have already noted that the assumption of a "rigid" bar is not consistent with the theory of relativity. What relativity means is that mechanical quantities obey the same transformation laws, whether their origin is mechanical or electrical. Hence forces derived from elastic stresses would depend on the velocity in the same way as those corresponding to the Lorentz force. The problem as a whole is similar to that in which the torque is balanced by the gain in angular momentum: in every case equilibrium is a property invariant under a Lorentz transformation. It is clear that to an observer moving along with the charges they do not actually constitute current elements, and the observed interaction will be just the static Coulomb force.

18–5 The virtual photon concept.

We have shown that the electromagnetic field of a rapidly moving charge approaches the field of a plane wave as the velocity approaches c. It is of considerable interest to express this correspondence in a quantitative way.

The electromagnetic field due to a passing charge represents a pulse in time which corresponds to a distribution in frequency of the energy contained in the field. Clearly, the integral of this frequency spectrum will simply represent the total energy of the field of the charge. This integral will, of course, be divergent unless the field is "cut off" at small distances.

Let us first make a Fourier analysis of the transverse electric field. We have, from Eq. (18–13),

$$\frac{4\pi\epsilon_0}{e} E_{\perp}(t) = \frac{\gamma^{-2}b}{[(ut)^2 + \gamma^{-2}b^2]^{3/2}}, \tag{18–30}$$

where $\gamma = (1 - u^2/c^2)^{-1/2}$ and b is the perpendicular distance between the particle trajectory and the observer (see Fig. 18–4). The Fourier components of this field are given by

$$\begin{aligned} E_{\omega\perp} &= \frac{1}{2\pi}\int_{-\infty}^{\infty} E_{\perp}(t)e^{i\omega t}\,dt \\ &= \frac{eb\gamma^{-2}}{8\pi^2\epsilon_0}\int_{-\infty}^{\infty}\frac{e^{i\omega t}}{[(ut)^2+\gamma^{-2}b^2]^{3/2}}\,dt \\ &= \frac{e}{8\pi^2\epsilon_0 bu}\int_{-\infty}^{\infty}\frac{e^{i(\omega b/\gamma u)\xi}}{(\xi^2+1)^{3/2}}\,d\xi, \end{aligned} \tag{18–31}$$

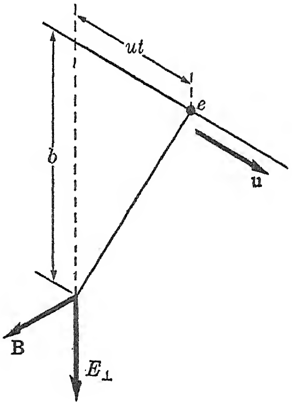


FIG. 18–4 For computing the transverse field of a moving electron.

where $\xi = \gamma ut/b$. This can be written as

$$E_{\omega\perp} = \frac{e}{4\pi^2\epsilon_0 bu}\left[\left(\frac{\omega b}{\gamma u}\right)K_1\left(\frac{\omega b}{\gamma u}\right)\right], \tag{18–32}$$

where K_1 is the Bessel function of the second kind with imaginary argument. The function $K_1(x)$ behaves like $1/x$ near $x = 0$ and goes to zero asymptotically as e^{-x}. Hence we can approximate Eq. (18–31) by

$$\begin{aligned} E_{\omega\perp} &= \frac{e}{4\pi^2\epsilon_0 bu}, \qquad \omega < \gamma u/b, \\ E_{\omega\perp} &= 0, \qquad \omega > \gamma u/b. \end{aligned} \tag{18–33}$$

This relation is quite reasonable, since due to the transverse contraction of the field the effective "time of passage" is of order $b/\gamma u$.

In an electromagnetic wave half the energy is carried by the electric and half by the magnetic field. Hence the total energy is

$$U = \epsilon_0\int E_{\perp}^2\,dv = \epsilon_0\int\left(\int E_{\perp}^2 u\,dt\right)2\pi b\,db. \tag{18–34}$$

By using the theorem proved in Eq. (13–36) we may write this energy in terms of Fourier components as

$$\begin{aligned} U &= \int_0^{\infty} U_\omega\,d\omega = 4\pi\epsilon_0 u\int\left(\int_0^{\infty}|E_{\omega\perp}|^2\,d\omega\right)2\pi b\,db \\ &= \frac{2}{\pi}\frac{e^2}{4\pi\epsilon_0 u}\int_0^{\infty}\int_{b_{\min}}^{\gamma u/\omega}\frac{db}{b}\,d\omega, \end{aligned} \tag{18–35}$$

where $b_{\min}$ is an arbitrary lower cutoff for the Coulomb field. Hence,

$$U_\omega = \frac{2}{\pi}\frac{e^2}{4\pi\epsilon_0 u}\ln\left(\frac{\gamma u}{\omega b_{\min}}\right) \tag{18–36}$$

AGRICULTURAL

gives the energy distribution of the equivalent plane wave field in frequency. Clearly, $b_{\min}$ remains undefined within the framework of classical theory.

It is attractive to translate Eq. (18–36) into the number of "equivalent photons" N_ω by the relation $U_\omega \, d\omega = \hbar\omega N_\omega \, d\omega$. Let us take $u \simeq c$. Then

$$N_\omega \, d\omega = (2\alpha/\pi) \ln (\gamma c/\omega b_{\min})(d\omega/\omega), \tag{18–37}$$

where $\alpha = e^2/4\pi\epsilon_0\hbar c \sim 1/137$ is the fine structure constant. Quantum mechanically the position of the charge is undefined within distances smaller than those of order $\hbar/m_0c$, so that we can take $b_{\min} \simeq A\hbar/m_0c$, where A is a numerical constant. Since $\gamma m_0 c^2$ is the total energy E of the particle, we can write

$$N_\omega \, d\omega = (2\alpha/\pi) \ln (E/A\hbar\omega) \, (d\omega/\omega). \tag{18–38}$$

Hence the spectrum of "equivalent photons" varies approximately as $1/\omega$. The number of equivalent photons per electron is small, namely, of the order of 1/137.

Equation (18–38) is very useful in relating the probability of processes induced by electrons, or by other particles which act essentially only through their electromagnetic field, to the probability of processes induced by electromagnetic radiation.

SUGGESTED REFERENCES

R. BECKER, *Theorie der Elektrizität*, Band II. Our treatment parallels that found in Section 11 of this work.

L. LANDAU AND E. LIFSHITZ, *The Classical Theory of Fields*. The Liénard-Wiechert potentials are derived from the four-dimensional tensor formulation.

A. SOMMERFELD, *Electrodynamics*. The differential equation for the four-potential is solved by integration in the complex plane.

J. A. STRATTON, *Electromagnetic Theory*. Another presentation of the complex integration of the wave equation. Sommerfeld has pointed out that the method was devised (by G. Herglotz) in 1904, before the theory of relativity, on the basis of mathematical elegance.

EXERCISES

1. Obtain the general expression for the field of a uniformly moving charge by making a Lorentz transformation of the static Coulomb field. How do you reconcile the statement that the field is decreased in the line of motion with Eq. (17–40), according to which $E'_{||} = E_{||}$?

2. Find the torque on the charges of Fig. 18–3 as derived from the convection potential and compare the answer with that found in Chapter 14.

CHAPTER 19

RADIATION FROM AN ACCELERATED CHARGE

We shall now consider the fields of a charge e in the general case for which the retarded position as a function of the retarded time is assumed given; in other words, $x'_\alpha(t')$ is known, where t' is the time at which a signal propagated with velocity c is emitted at x'_α so as to arrive at x_α at time t. Note that this statement implies the existence of a definite means for establishing the relation between the time of emission, t', and the time t of reception at the field point. Prior to relativity theory, it was simply assumed that the problem is solved in the frame for which the wave equation involving c is valid, and hence t and t' are connected by c. Relativity extends the validity of this formulation to all frames. The additional difference between the classical and relativistic formulation is conceptual. In classical physics it is assumed that there exists a universal time scale, so that the connection between t' and t could be verified independently; hence the validity of the assumption that a frame is, in fact, the one for which the wave equation is correct could be independently verified. Relativity denies this hypothesis. Moreover, we shall have to consider the meaning of the velocity $\mathbf{u}$ in the two cases. In classical theory $\mathbf{u}$ is the velocity of the source point in the frame in which the wave equation is valid; in special relativity $\mathbf{u}$ is the velocity in any inertial frame.

19–1 Fields of an accelerated charge. Let us compute the complete electric and magnetic fields of a charge e for which $x'_\alpha(t')$ is given. The retarded values of the velocity and acceleration of the charge,

$$u_\alpha = \partial x'_\alpha/\partial t', \quad \dot{u}_\alpha = \partial^2 x'_\alpha/\partial t'^2, \tag{19–1}$$

are thus also known. Vectorially,

$$\frac{d\mathbf{r}}{dt'} = -\mathbf{u}, \quad \frac{d^2\mathbf{r}}{dt'^2} = \dot{\mathbf{u}}. \tag{19–2}$$

The Liénard-Wiechert potentials are given by the usual formulas,

$$\phi(x_\alpha,t) = \frac{e}{4\pi\epsilon_0}\frac{1}{s}, \tag{18–6}$$

$$\mathbf{A}(x_\alpha,t) = \frac{e}{4\pi\epsilon_0 c^2}\frac{\mathbf{u}}{s}, \tag{18–7}$$

where $s = r - (\mathbf{u} \cdot \mathbf{r})/c$ is a function of both the field point and the retarded source point coordinates. The field and source point variables are connected by the retardation condition:

$$r(x_\alpha, x'_\alpha) = [\Sigma(x_\alpha - x'_\alpha)^2]^{1/2} = c(t - t'). \tag{19-3}$$

When we derive the fields $\mathbf{E}$ and $\mathbf{B}$ from the potentials of Eqs. (18-6) and (18-7) in the usual way,

$$\mathbf{B} = \nabla \times \mathbf{A},$$

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t},$$

we must notice that the components of the vector operator are partial derivatives at constant time t, and therefore *not* at constant time t'. Partial differentiation with respect to x_α compares the potentials at neighboring points at the same time, but these potential signals originated from the charge at different times. Similarly, the partial with respect to t implies constant x_α, and hence refers to the comparison of potentials at a given field point over an interval of time during which the coordinates of the source will have changed. Since only the time variation with respect to t' is given (in the original description of the problem) we must transform $\left.\frac{\partial}{\partial t}\right|_{x_\alpha}$ and $\nabla|_t$ to expressions in terms of $\left.\frac{\partial}{\partial t'}\right|_{x_\alpha}$ in order to compute the fields. This procedure is necessary because it is, in general, impossible in the case of an accelerated source charge to express the potentials in terms of the "present position" alone, as we did in the case of uniform motion.

To obtain the required transformation for the derivatives, we note that since x'_α is assumed given as a function of t', we have from Eq. (19-3),

$$r[x_\alpha, x'_\alpha(t')] = f(x_\alpha, t') = c(t - t'), \tag{19-4}$$

which is a functional relation between x_α, t, and t'. In order to relate $\partial/\partial t$ to $\partial/\partial t'$, we need only hold x_α fixed in this relation and note that

$$\left(\frac{\partial r}{\partial t'}\right)_{x\alpha} = -\frac{\mathbf{r} \cdot \mathbf{u}}{r}. \tag{19-5}$$

Then

$$\frac{\partial r}{\partial t} = c\left(1 - \frac{\partial t'}{\partial t}\right) = \frac{\partial r}{\partial t'}\frac{\partial t'}{\partial t} = -\frac{\mathbf{r} \cdot \mathbf{u}}{r}\frac{\partial t'}{\partial t}, \tag{19-6}$$

or

$$\frac{\partial t'}{\partial t} = \frac{1}{1 - \dfrac{\mathbf{r} \cdot \mathbf{u}}{rc}} = \frac{r}{s}. \tag{19-7}$$

Therefore,

$$\frac{\partial}{\partial t} = \frac{r}{s}\frac{\partial}{\partial t'} \tag{19-8}$$

is the desired transformation for the time derivatives. Similarly, for the vector operator ∇,

$$\nabla r = -c\nabla t' = \nabla_1 r + \frac{\partial r}{\partial t'}\nabla t' = \frac{\mathbf{r}}{r} - \frac{\mathbf{r}\cdot\mathbf{u}}{r}\nabla t', \tag{19–9}$$

where by ∇_1 we mean differentiation with respect to the first argument of the function f in Eq. (19–4), that is, differentiation at constant retarded time t'. Therefore, from Eq. (19–9),

$$\nabla t' = -\frac{\mathbf{r}}{sc}, \tag{19–10}$$

and, in general

$$\nabla = \nabla_1 - \frac{\mathbf{r}}{sc}\frac{\partial}{\partial t'}. \tag{19–11}$$

Equations (19–8) and (19–11) constitute the required transformation of the differential operators from the coordinates of the field point to those of the radiator.

The computation of the electric field from the Liénard-Wiechert potentials thus becomes

$$\begin{aligned}\frac{4\pi\epsilon_0}{e}\mathbf{E} &= \frac{1}{s^2}\nabla s - \frac{\partial}{\partial t}\frac{\mathbf{u}}{sc^2}\\ &= \frac{1}{s^2}\nabla_1 s - \frac{\mathbf{r}}{cs^3}\frac{\partial s}{\partial t'} - \frac{r}{s^2c^2}\dot{\mathbf{u}} + \frac{r\mathbf{u}}{c^2s^3}\frac{\partial s}{\partial t'}.\end{aligned} \tag{19–12}$$

Using $\nabla_1 s = \mathbf{r}/r - \mathbf{u}/c$ and Eq. (19–5), and then collecting terms, we find that the two terms containing $\dot{\mathbf{u}}$ combine to a vector triple product, so that

$$\frac{4\pi\epsilon_0}{e}\mathbf{E} = \frac{1}{s^3}\left(\mathbf{r} - \frac{r\mathbf{u}}{c}\right)\left(1 - \frac{u^2}{c^2}\right) + \frac{1}{c^2s^3}\left\{\mathbf{r}\times\left[\left(\mathbf{r} - \frac{r\mathbf{u}}{c}\right)\times\dot{\mathbf{u}}\right]\right\}. \tag{19–13}$$

Similarly,

$$\begin{aligned}\frac{4\pi\epsilon_0c^2}{e}\mathbf{B} = \nabla\times\frac{\mathbf{u}}{s} &= -\frac{\mathbf{r}\times\dot{\mathbf{u}}}{cs^2} + \frac{\mathbf{u}\times\mathbf{r}}{s^2}\left[\frac{1}{r} + \frac{1}{s}\left(\frac{\mathbf{r}\cdot\mathbf{u}}{rc} + \frac{\mathbf{r}\cdot\dot{\mathbf{u}}}{c^2} - \frac{u^2}{c^2}\right)\right]\\ &= -\frac{\mathbf{r}\times\dot{\mathbf{u}}}{cs^2} + \frac{\mathbf{u}\times\mathbf{r}}{s^3}\left(1 - \frac{u^2}{c^2} + \frac{\mathbf{r}\cdot\dot{\mathbf{u}}}{c^2}\right)\\ &= \frac{\mathbf{u}\times\mathbf{r}}{s^3}\left(1 - \frac{u^2}{c^2}\right) + \frac{1}{cs^3}\frac{\mathbf{r}}{r}\times\left\{\mathbf{r}\times\left[\left(\mathbf{r} - \frac{r\mathbf{u}}{c}\right)\times\dot{\mathbf{u}}\right]\right\}.\end{aligned} \tag{19–14}$$

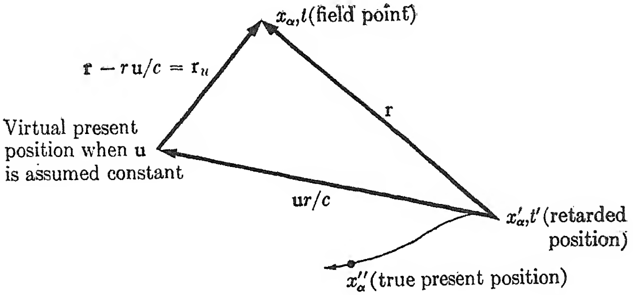

x_α, t (field point)
$\mathbf{r} - r\mathbf{u}/c = \mathbf{r}_u$
$\mathbf{r}$
Virtual present position when $\mathbf{u}$ is assumed constant
$\mathbf{u}r/c$
x'_α, t' (retarded position)
x''_α (true present position)

FIG. 19-1 Position parameters for the field of a charge in arbitrary motion.

On comparison with Eq. (19-13), we see that

$$\mathbf{B} = \frac{\mathbf{r} \times \mathbf{E}}{rc}. \tag{19-15}$$

Thus the magnetic field is always perpendicular to $\mathbf{E}$ and to the *retarded* radius vector $\mathbf{r}$.

The electric field is composed of two separate parts. The first term in Eq. (19-13) varies as $1/r^2$ for large distances, and is formally identical with the "convective" field of a uniformly moving charge. We might call $\mathbf{r}_u = \mathbf{r} - r\mathbf{u}/c$ the "virtual present radius vector," i.e., the position the charge would occupy "at present" if it had continued with uniform velocity from the point x'_α (see Fig. 19-1). In terms of $\mathbf{r}_u$ the $1/r^2$ field is simply

$$\mathbf{E}_{\text{induction}} = \frac{e\mathbf{r}_u}{4\pi\epsilon_0 s^3}\left(1 - \frac{u^2}{c^2}\right), \tag{19-16}$$

which is identically Eq. (18-14) for a charge in uniform motion. It is analogous to the quasi-static or induction field which we discussed in connection with the radiation from variable current and charge systems in Chapter 13. Equation (19-16) represents a nonradiating term in the sense that it does not contribute to the energy flow over an infinitely distant surface, but it may contribute to the net energy loss of the electron. If the charge is accelerated, the field is neither static nor convective, and there is a net change in field energy which must be supplied. This energy loss will cause a reaction on whatever outside force is responsible for the electron's motion. We shall return to a detailed consideration of this reaction force in the next chapter.

The second term in Eq. (19-13),

$$\mathbf{E}_{\text{radiation}} = \frac{e}{4\pi\epsilon_0 s^3 c^2}\,\mathbf{r} \times (\mathbf{r}_u \times \dot{\mathbf{u}}), \tag{19-17}$$

is of order $1/r$ and therefore does represent a radiation field in the sense of contributing to the energy flux over a large sphere. Similar considerations hold for the two terms of Eq. (19–14). Let us look at some important special cases of this radiation field.

19–2 Radiation at low velocity. In case the velocity is so small that u/c is negligible in comparison with unity, $\mathbf{r}_u \simeq \mathbf{r}$ and $s \simeq r$, so that

$$\mathbf{E}_{\text{rad}} = \frac{e}{4\pi\epsilon_0 c^2 r^3}\,\mathbf{r} \times (\mathbf{r} \times \dot{\mathbf{u}}), \tag{19–18a}$$

$$\mathbf{B}_{\text{rad}} = \frac{e}{4\pi\epsilon_0 c^3 r^2}\,\dot{\mathbf{u}} \times \mathbf{r}. \tag{19–18b}$$

Equations (19–18a, b) represent a field which is formally identical to that [Eq. (13–59)] of a radiating electric dipole of moment equal to $\dot{\mathbf{u}}e/\omega^2$. The angular distribution of the radiated energy is therefore simply the $\sin^2\theta$ distribution discussed in Section 13–6 (see Fig. 19–2). To obtain the rate of energy loss from the accelerated charge, we integrate over a sphere the Poynting vector corresponding to Eqs. (19–18) and find

$$-\frac{dW}{dt'} = \frac{e^2\dot{u}^2}{6\pi\epsilon_0 c^3}. \tag{19–19}^*$$

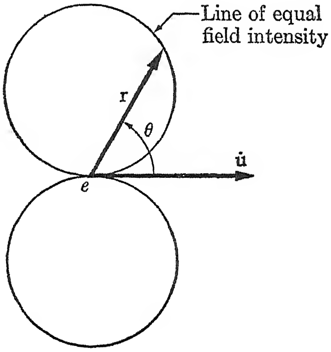


FIG. 19–2 Angular distribution of radiation field of a low velocity electron. The pattern is symmetrical about the axis of θ.

19–3 The case of $\dot{u}$ parallel to u. If $\dot{\mathbf{u}}$ and $\mathbf{u}$ are along the same direction, regardless of whether u is small, the radiation fields are simply

$$\mathbf{E} = \frac{e}{4\pi\epsilon_0 c^2 s^3}\,\mathbf{r} \times (\mathbf{r} \times \dot{\mathbf{u}}), \tag{19–20a}$$

$$\mathbf{B} = \frac{er}{4\pi\epsilon_0 c^3 s^3}\,\dot{\mathbf{u}} \times \mathbf{r}. \tag{19–20b}$$

---

* In Gaussian units (see Appendix I),

$$\frac{-dW}{dt'} = \frac{2e^2\dot{u}^2}{3c^3}.$$

All equations for radiation rates in this and the succeeding chapters are written so that they can be reduced to Gaussian units by putting $\left(\frac{e^2}{4\pi\epsilon_0}\right)_{\text{mks}} = (e^2)_{\text{Gaussian}}$.

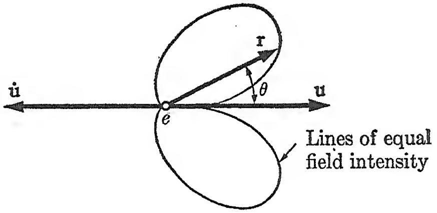


FIG. 19–3 Sketch of the radiation field of an electron decelerated along its line of motion, **u** comparable to c.

Equations (19–20) differ only by the factor $r^3/s^3 = [1 - (u/c)\cos\theta]^{-3}$ from the slow electron (dipole) case of Eqs. (19–18). The qualitative effect of this factor is to increase the radiated energy in the forward direction, as shown in Fig. 19–3.

To calculate quantitatively the angular distribution of the radiated energy, we must look carefully at what is meant by "the rate of radiation" of the charge. This is the amount of energy lost by the electron in a time interval dt' during the emission of the signal, i.e., the rate of energy loss $-dW/dt'$ of the electron itself. At a given field point the Poynting vector **N** represents the energy flow per unit time measured as t. Therefore the rate of energy loss of the electron into a given solid angle $d\Omega$ is given by

$$-\frac{dW(\theta)}{dt'}\,d\Omega = |\mathbf{N}|\,\frac{dt}{dt'}\,r^2\,d\Omega = |\mathbf{E}\times\mathbf{H}|\,\frac{dt}{dt'}\,r^2\,d\Omega$$

$$= \epsilon_0 c E^2 r^2\,\frac{dt}{dt'}\,d\Omega = \epsilon_0 c E^2\,\frac{s}{r}\,r^2\,d\Omega, \qquad (19\text{–}21)$$

when we have used Eq. (19–7). Hence the directional rate of energy loss for the accelerated electron is

$$-\frac{dW(\theta)}{dt'} = \frac{\dot{u}^2}{c^3}\left(\frac{e^2}{16\pi^2\epsilon_0}\right)\frac{\sin^2\theta}{\left(1 - \dfrac{u}{c}\cos\theta\right)^5} \qquad (19\text{–}22)$$

The correction s/r can be physically interpreted as follows. The energy emitted by the electron in a time dt' is located in the volume between two spheres, one of radius r about a point $(x'_\alpha)_2$ and the other of radius $r + c\,dt'$ about a point $(x'_\alpha)_1$ which is a distance $-u\,dt'$ from the second position of the source, as indicated in Fig. 19–4. Consider the element dv of this volume subtending a solid angle $d\Omega = dS/r^2$ at $(x'_\alpha)_2$. Since $dr = c\,dt' - [(\mathbf{r}\cdot\mathbf{u})/r]\,dt'$,

$$dv = dS\,dr = dS\left(c - \frac{\mathbf{u}\cdot\mathbf{r}}{r}\right)dt' = \frac{cs}{r}\,dS\,dt'. \qquad (19\text{–}23)$$

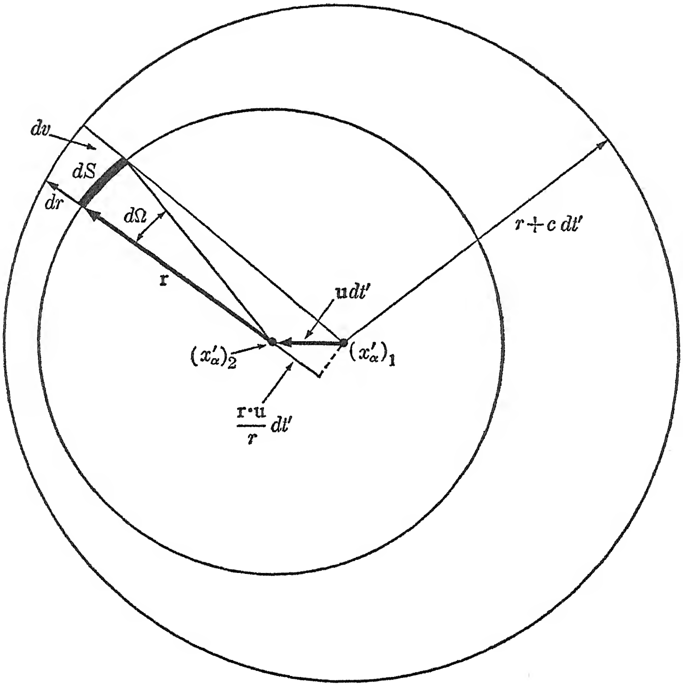


FIG. 19–4 Location of energy radiated by an electron as it moves from $(x'_\alpha)_1$ to $(x'_\alpha)_2$.

Therefore the energy $dW\,d\Omega$ contained in this volume within the solid angle $d\Omega$ is

$$dW\,d\Omega = \frac{\epsilon_0 E^2 + \mu_0 H^2}{2}\,\frac{cs}{r}\,dS\,dt' = \epsilon_0 E^2\,\frac{cs}{r}\,dS\,dt' \qquad (19\text{–}24)$$

in agreement with Eq. (19–21).

The principal application of Eqs. (19–20) and (19–21) is to the calculation of radiation from an electron which is decelerated, it being assumed in this approximation that the direction of motion does not change. Deceleration radiation is known as *bremsstrahlung*. For an exact classical calculation it would be necessary to put into the equations the precise variation of acceleration with time, using the stopping power of the material upon which the electron impinges. For a simplified discussion, let us assume that $\dot{u}$ is constant while the velocity decreases from u_0 to 0. This gives a resultant pulse of radiated energy:

$$dW = \frac{-e^2 \sin^2\theta\,\dot{u}}{(4\pi)^2 \epsilon_0 c^3} \int_{u_0}^{0} \frac{\dot{u}\,dt}{\left(1 - \frac{u}{c}\cos\theta\right)^5}\,d\Omega, \qquad (19\text{–}25)$$

$$-dW = \frac{e^2 \sin^2 \theta \dot{u}}{64\pi^2 \epsilon_0 c^2 \cos\theta}\left[\frac{1}{\left(1 - \frac{u_0}{c}\cos\theta\right)^4} - 1\right] d\Omega. \tag{19-25}$$

Thus the angular distribution of the pulse as a whole is also tipped forward in the direction of the motion. The radiation is polarized with the electric vector lying in the plane of the radius vector and the direction of deceleration. Equation (19–25) can be used to estimate the total efficiency of an x-ray tube. In practice, however, both the angular distribution and the polarization of the outgoing radiation are greatly modified by scattering of the electrons in the target material.

The frequency spectrum of the outgoing radiation can be obtained by Fourier analysis of the radiation field. Let us assume for simplicity that a change Δu in the velocity takes place in a very short time $\Delta t'$, and that $u \ll c$. If the radiation takes place at time t_0, the wave field during the short time interval may be written

$$E(t) = \frac{e \sin\theta}{4\pi\epsilon_0 c^2 r}\frac{\Delta u}{\Delta t'} = \frac{e\sin\theta}{4\pi\epsilon_0 c^2 r}\Delta u\, \delta(t - t_0), \tag{19-26}$$

where we have expressed $\dot{u}$ as a δ-function,

$$\dot{u} = \delta(t - t_0)\,\Delta u, \qquad \int_{-\infty}^{\infty} \dot{u}\, dt = \Delta u.$$

If we put

$$E(t) = \int_{-\infty}^{\infty} E_\omega e^{-i\omega t}\, d\omega,$$

then

$$E_\omega = \frac{1}{2\pi}\int E(t) e^{i\omega t}\, dt = \frac{e\sin\theta\,\Delta u}{8\pi^2\epsilon_0 c^2 r} e^{i\omega t_0}, \tag{19-27}$$

which, except for phase, is independent of ω. Now the total energy loss is obtained by integrating the Poynting vector over the surface of a sphere and over the time during which the change in velocity takes place. For $u \ll c$,

$$-W = \int -\frac{dW}{dt'}\, dt' = \epsilon_0 c \int\int_{-\infty}^{\infty} E^2\, dt\, dS. \tag{19-28}$$

The frequency of the energy loss, corresponding to a field whose Fourier components are E_ω, is then obtained as in Section 13–3.

$$\int_{-\infty}^{\infty} E^2\, dt = 4\pi \int_0^{\infty} |E_\omega|^2\, d\omega \tag{19-29}$$

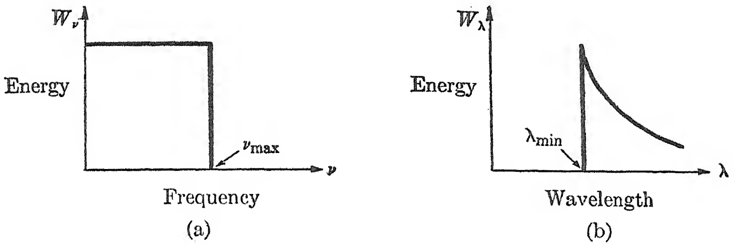


FIG. 19-5 Ideal spectrum of bremsstrahlung as a function of frequency and as a function of wavelength.

is analogous to Eq. (13-36). Hence the energy loss in a frequency band $d\omega$ is

$$-W_\omega\,d\omega = 4\pi\epsilon_0 c\int_S |E_\omega|^2\,dS\,d\omega. \tag{19-30}$$

For our specific spectrum, Eq. (19-27),

$$\begin{aligned}-W_\omega\,d\omega &= \frac{e^2}{4\pi\epsilon_0 c}\left(\frac{\Delta u}{c}\right)^2\frac{d\omega}{(2\pi)^2}\int_0^{2\pi}d\varphi\int_0^{\pi}\sin^2\theta\sin\theta\,d\theta\\ &= \frac{e^2}{3\pi\epsilon_0 c}\left(\frac{\Delta u}{c}\right)^2\frac{d\omega}{2\pi}.\end{aligned} \tag{19-31}$$

Equation (19-31) indicates that the spectrum of radiation is constant on a frequency scale. Even classically the extent of the spectrum to high ω is due only to the simplifying assumption of zero collision time: a Fourier analysis of a finite collision time process will automatically remove the very high frequency components. Actually the spectrum will be cut off at the point where the kinetic energy of the electron is equivalent to a single quantum of radiation:

$$\text{K.E.} = \tfrac{1}{2}mu^2 = h\nu_{\max}. \tag{19-32}$$

With such a maximum the ideal spectrum is shown on a frequency scale and again on a wavelength scale in Fig. 19-5. By means of Planck's hypothesis, Eq. (19-31) can be expressed in terms of the number of quanta, $dN = -W_\omega\,d\omega/\hbar\omega$, that are "shaken off" during the velocity change. This gives

$$dN_\omega = \frac{e^2}{4\pi\epsilon_0\hbar c}\frac{2}{3\pi}\left(\frac{\Delta u}{c}\right)^2\frac{d\omega}{\omega} = \frac{1}{137}\frac{2}{3\pi}\left(\frac{\Delta u}{c}\right)^2\frac{d\omega}{\omega}. \tag{19-33}$$

Thus an infinite number of zero energy quanta would be emitted, although the total energy is finite. This feature is also present in a more exact treatment, and is known as the *infrared catastrophe.*

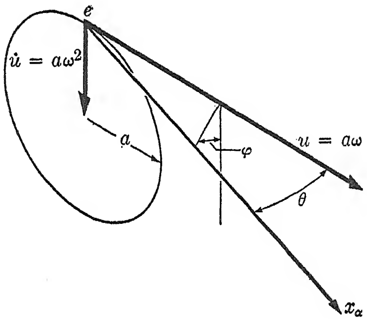


FIG. 19–6 Charge with acceleration at right angles to its velocity.

19–4 Radiation when the acceleration is perpendicular to the velocity. There are important practical applications of the case of a charge moving in a circle of radius a with a constant angular velocity ω, as indicated in Fig. 19–6. Here $u = a\omega$, $\dot{u} = a\omega^2$, $\mathbf{r}_u = \mathbf{r} - \mathbf{u}r/c$, $\mathbf{u} \cdot \mathbf{r} = ur\cos\theta$, $\dot{\mathbf{u}} \cdot \mathbf{r} = \dot{u}r\sin\theta\cos\varphi$, where φ is the azimuthal angle. Hence

$$[\mathbf{r} \times (\mathbf{r}_u \times \dot{\mathbf{u}})]^2 = \dot{u}^2 r^4 \left[\left(1 - \frac{u}{c}\cos\theta\right)^2 - \left(1 - \frac{u^2}{c^2}\right)\sin^2\theta\cos^2\varphi\right]$$

and thus

$$-\frac{dW(\theta,\varphi)}{dt'}\,d\Omega = \frac{e^2\dot{u}^2}{16\pi^2\epsilon_0 c^3}\,\frac{\left(1 - \dfrac{u}{c}\cos\theta\right)^2 - \left(1 - \dfrac{u^2}{c^2}\right)\sin^2\theta\cos^2\varphi}{\left(1 - \dfrac{u}{c}\cos\theta\right)^5}\,d\Omega. \tag{19–34}$$

The resultant pattern has zeros in the plane of the circle at $\theta = \cos^{-1}(u/c)$. For large velocities the radiation is very much more intense in the forward direction than in any other, as indicated in Fig. 19–7. As u approaches c the radiation becomes a sharp forward ray.

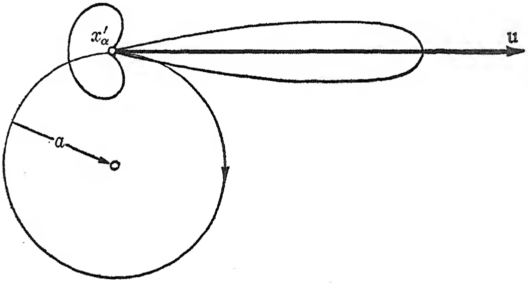


FIG. 19–7 Sketch of the radiation field of Fig. 19–6 in the plane of motion for **u** comparable with c. The pattern is not symmetrical about the axis of θ in this case.

Integration of Eq. (19–34) gives the total rate of radiation of a charge e moving in a circle of radius a with constant angular velocity ω:

$$-\frac{dW}{dt'} = \frac{e^2\dot{u}^2}{6\pi\epsilon_0 c^3}\frac{1}{(1-u^2/c^2)^2} = \frac{e^2a^2\omega^4}{6\pi\epsilon_0 c^3}\left(\frac{W}{W_0}\right)^4, \tag{19–35}$$

where $W_0 = m_0c^2$ and W is the relativistic energy $W = W_0/\sqrt{1-u^2/c^2}$. Such considerations are important in the theory of the betatron and synchrotron. The betatron, for example, is a machine in which electrons are accelerated to very high energies while moving in a circle of fixed radius. Some numerical examples of the energy which must be supplied for radiation loss are included in the problems.

19–5 Radiation with no restrictions on the acceleration or velocity.

The general radiation field, Eq. (19–17), together with the correction to the rate of radiation that leads to Eq. (19–21), gives the general expression for the directional rate of radiation:

$$-\frac{dW}{dt'}\,d\Omega = \frac{e^2r}{16\pi^2\epsilon_0 s^5c^3}[\mathbf{r}\times(\mathbf{r}_u\times\dot{\mathbf{u}})]^2\,d\Omega. \tag{19–36}$$

Since the radiation fields vanish when $\mathbf{r}_u = \mathbf{r} - \mathbf{u}r/c$ is parallel to $\dot{\mathbf{u}}$, there are, in general, two nodal lines such as we have noted in the special cases. The position of the nodes can be constructed graphically, as shown in Fig. 19–8. Construct a circle of radius r about the charge at O, and lay off $OQ = \mathbf{u}r/c$. A line through Q parallel to $\dot{\mathbf{u}}$ with end points A and B on the circle will then represent both values of $\mathbf{r}_u$ for which the fields vanish. OA and OB are thus the two nodal lines, which always lie in the plane of $\mathbf{u}$ and $\dot{\mathbf{u}}$.

It is an elementary but somewhat complicated matter to integrate Eq. (19–36) over the total solid angle. Probably the simplest method is to

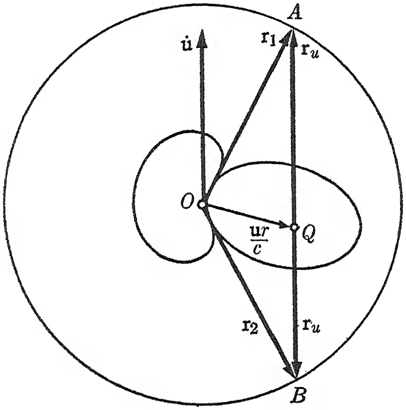


FIG. 19–8 Graphical construction of nodal lines in the radiation field.

choose the direction of $\mathbf{u}$ as the polar axis and first perform the azimuthal integration. The final result of the integration is

$$-\frac{dW}{dt'} = \frac{e^2}{6\pi\epsilon_0 c^3} \frac{\dot{u}^2 - (\mathbf{u} \times \dot{\mathbf{u}})^2/c^2}{(1 - u^2/c^2)^3} \tag{19–37}$$

for the total rate of radiation. It is easy to check that the radiation rates given in the previous sections are special cases of Eq. (19–37). That Eq. (19–37) can be derived by Lorentz transformation from Eq. (19–19) is shown in the problems.

19–6 Classical cross section for bremsstrahlung in a Coulomb field. At low electron energies Eq. (19–25) gives the approximate yield of "soft" x-rays if empirical values of $\dot{u}$ are used. In that case, collisions of the primary electrons with nuclei are rare, and the principal source of energy loss and momentum transfer consists of electron-electron collisions. But single collisions between particles of like e/m do not produce either electric dipole or magnetic dipole radiation, since displacement and circulation of the center of charge are inconsistent with the conservation of linear and angular momentum. (As an example, see problem 4 in the exercises at the end of this chapter.) Hence, in the low-energy range, where electron-electron collisions predominate, the radiation is computed by the collective effects of the electrons in retarding the incident particle. At higher energies, and particularly for heavy targets, the x-ray yield must be considered as being due to radiation during the deflection of the electron in the Coulomb field of a nucleus. For electron collisions with nuclei, dipole radiation is possible, and in these circumstances the effect of single collisions predominates. We can make a simple nonrelativistic calculation, involving only the transverse acceleration, which contains the essential physical features of a more detailed treatment.

Consider an electron of velocity u passing a nucleus of charge Ze at a distance b. If the time of closest approach is taken as $t = 0$, the magnitude of the force as a function of the time is

$$F = m\dot{u} = \frac{Ze^2}{4\pi\epsilon_0(b^2 + u^2t^2)}, \tag{19–38}$$

so that the (small) transverse component of acceleration suffered by the electron is approximately

$$\dot{u} = \frac{Ze^2}{4\pi\epsilon_0 m} \frac{b}{(b^2 + u^2t^2)^{3/2}}. \tag{19–39}$$

To calculate the radiation loss and the spectral distribution, we must make a Fourier analysis of $\dot{u}$:

$$\dot{u}_\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} \dot{u}(t)e^{-i\omega t}\,dt = \frac{1}{2\pi}\frac{2Ze^2b}{4\pi\epsilon_0 m}\int_0^{\infty} \frac{\cos\omega t\,dt}{(b^2 + u^2t^2)^{3/2}}. \tag{19–40}$$

Now the Bessel function $K_1(\omega b/v)$ is represented by the integral

$$K_1\left(\frac{\omega b}{u}\right) = \frac{b}{u\omega}\int_0^\infty \frac{\cos \omega t\, dt}{[(b/u)^2 + t^2]^{3/2}}, \tag{19-41}$$

and thus

$$\dot{u}_\omega = \frac{1}{2\pi}\frac{Ze^2}{4\pi\epsilon_0 m}\frac{2\omega}{u^2} K_1\left(\frac{\omega b}{u}\right). \tag{19-42}$$

The function $K_1(\omega b/u)$ has the property that $K_1(\omega b/u) \simeq u/\omega b$ for $\omega b/u < 1$, and for large values of the argument it decreases exponentially. We may therefore take

$$\begin{aligned} \dot{u}_\omega &= \frac{1}{\pi}\frac{Ze^2}{4\pi\epsilon_0 m}\frac{1}{ub}, \quad \omega < u/b, \\ \dot{u}_\omega &= 0, \quad \omega > u/b. \end{aligned} \tag{19-43}$$

The radiated energy per unit solid angle is then, from Eq. (19–30) with the substitution of the radiation of the radiation field appropriate to $\dot{u}_\omega$,

$$-W_\omega\, d\omega\, d\Omega = \frac{Z^2e^2r_0^2}{4\pi\epsilon_0}\frac{1}{\pi^2}\frac{c}{u^2b^2}\sin^2\theta\, d\omega\, d\Omega, \tag{19-44}$$

where $r_0 = e^2/4\pi\epsilon_0 mc^2$ is the classical electron radius. The total rate of radiation per frequency interval $d\omega$ is

$$-W_\omega\, d\omega = \frac{2Z^2e^2r_0^2}{3\pi^2\epsilon_0}\frac{c}{u^2b^2}\, d\omega. \tag{19-45}$$

Equation (19–45) can be expressed as a cross section for the emission of quanta of energy $\hbar\omega$, if we take account of all possible distances b:

$$\begin{aligned} d\sigma &= \int_{b_{\min}}^{b_{\max}} \frac{W_\omega\, d\omega}{\hbar\omega} 2\pi b\, db = \frac{4Z^2r_0^2}{3\pi\epsilon_0}\frac{e^2}{\hbar c}\frac{c^2}{u^2}\ln\left(\frac{b_{\max}}{b_{\min}}\right)\frac{d\omega}{\omega} \\ &= \frac{16\alpha}{3}\frac{Z^2r_0^2}{u^2/c^2}\ln\left(\frac{b_{\max}}{b_{\min}}\right)\frac{d\omega}{\omega}, \end{aligned} \tag{19-46}$$

in which it must be remembered that $\omega < u/b_{\min}$. The limit $b_{\max}$ is defined by the effective limit of the Coulomb field due to "screening" by atomic electrons surrounding the nucleus. The limit $b_{\min}$ is equivalent to cutting off the spectrum for $\hbar\omega$ greater than the kinetic energy of the incoming electron. Equation (19–46) is only approximate, but it illustrates the dependence on Z, r_0, and $\alpha = e^2/4\pi\epsilon_0\hbar c$ which is retained in the quantum mechanical calculation.

19–7 Čerenkov radiation. We have seen in Chapter 18 and again in Section 19–1 that a charge moving uniformly *in vacuo* does not radiate. The Poynting vector corresponding to the induction fields, Eq. (19–16) and the related magnetic field given by Eq. (19–15), falls off too rapidly with increasing r to contribute to a surface integral at large distances. The situation may be different, however, in a dielectric medium for which $\epsilon > \epsilon_0$, $\kappa > 1$. Let us consider the field of a moving charge in such a nonconducting medium, where for simplicity we let $\mu = \mu_0$. With the more general definition of the D'Alembertian operator given by Eq. (13–6), the inhomogeneous equation for the vector potential is just

$$\Box\mathbf{A} = \nabla^2\mathbf{A} - \frac{n^2}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} = -\mu_0\mathbf{j}. \tag{13–7}$$

The solutions are the same as the vacuum solutions except that c is replaced by c/n, where the index of refraction $n = \sqrt{\epsilon/\epsilon_0} = \sqrt{\kappa}$, and thus the retarded time is $t' = t - rn/c$. The relativistic limitation on the velocity of a particle is $u < c$, but if an electron with velocity nearly equal to c enters a dielectric medium its velocity may be greater than that of light in the medium. In these circumstances the retardation denominator may vanish, and in the direction for which $1 - un/c \cos\theta = 0$ the field intensities are infinite. In a cone defined by the angle $\theta = \cos^{-1}(c/nu)$ with the direction of motion, we might therefore expect to find radiation. Such radiation was first observed by Čerenkov in 1934. Note that if $u > c/n$ the transformation of Eq. (13–7) to an equivalent static solution, as discussed in Section 18–2, is impossible.

For a quantitative estimate of Čerenkov radiation it is convenient to use the Fourier solutions of Chapter 13. The current density $\mathbf{j}$ corresponding to a charge e moving with velocity $\mathbf{u}$ can be represented by

$$\mathbf{j} = e\mathbf{u}\,\delta(x' - ut)\,\delta(y')\,\delta(z'). \tag{19–47}$$

This expression evidently localizes the charge correctly at the source point (x',y',z'), and makes $\int \mathbf{j}\,dy'\,dz'\,dt = e\hat{\mathbf{x}}$, where $\hat{\mathbf{x}}$ is a unit vector in the direction of motion. The Fourier transform of $\mathbf{j}$ required for substitution into Eq. (13–33) is given by

$$\mathbf{j}_\omega = \frac{1}{2\pi}\int \mathbf{j}e^{i\omega t}\,dt = \frac{e}{2\pi}\,\delta(y')\,\delta(z')\,e^{i\omega x'/u}\hat{\mathbf{x}}. \tag{19–48}$$

Hence, if we let x', y', z' be the components of the source vector $\boldsymbol{\xi}$ in Eq. (13–33), the Fourier component of the magnetic field becomes

$$H_\omega = -\frac{i}{4\pi}\frac{e^{ikR}}{R}\frac{ek\sin\theta}{2\pi}\int e^{i\left(\frac{\omega x'}{u} - kx'\cos\theta\right)}\,dx' \tag{19–49}$$

when the integrals over y' and z' have been performed. Here $k = \omega n/c$ as defined in Eq. (11–19), and θ is the angle between $\mathbf{k}$ and $\mathbf{u}$. We can now follow the line of derivation leading to Eq. (13–37), with the slight modification due to the presence of the refractive index n in the equation.

In the wave zone the relation between the magnitudes of **E** and **H** is

$$|\mathbf{E}| = \frac{1}{n}\sqrt{\frac{\mu_0}{\epsilon_0}}\,|\mathbf{H}|,$$

and hence the energy loss in a frequency band $d\omega$ is

$$\begin{aligned}\delta W &= \frac{4\pi}{\epsilon_0 cn}\int |H_\omega|^2 R^2\, d\Omega\, d\omega \\ &= \frac{4\pi}{\epsilon_0 cn}\frac{e^2n^2\omega^2}{64\pi^4c^2}\int\left|\int_{-\infty}^{\infty} e^{i\left(\frac{\omega x'}{u} - kx'\cos\theta\right)}\, dx'\right|^2 \sin^2\theta\, d\Omega\, d\omega. \qquad (19\text{–}50)\end{aligned}$$

The integral over x' is essentially a δ-function, $\delta(1 - nu/c\cos\theta)$, again indicating infinite fields in the direction $\theta = \cos^{-1}(c/nu)$. But this form is easily manageable: an infinite path is not available to the electron, and the integral can be taken from $-X$ to X:

$$\int_{-X}^{X} e^{i\omega\left(1 - \frac{nu}{c}\cos\theta\right)x'/u}\, dx' = \frac{2\sin\left[\omega\left(1 - \dfrac{nu}{c}\cos\theta\right)\dfrac{X}{u}\right]}{\omega\left(1 - \dfrac{nu}{c}\cos\theta\right)\Big/ u}. \qquad (19\text{–}51)$$

Therefore

$$\delta W = \frac{1}{2\pi^2}\frac{e^2n\omega^2}{\epsilon_0c^3}\int_{-1}^{+1}\sin^2\theta\,\frac{\sin^2\left[\left(1 - \dfrac{nu}{c}\cos\theta\right)\dfrac{\omega X}{u}\right]}{\omega^2\left(1 - \dfrac{nu}{c}\cos\theta\right)^2\Big/ u^2}\, d(\cos\theta)\, d\omega. \qquad (19\text{–}52)$$

The angular integral still has some of the characteristics of a δ-function: it has a sharp maximum at $\cos\theta = c/nu$. If we use this value of θ in the slowly varying function $\sin^2\theta$, and remember that the limits do not matter so long as they include the one maximum, we may write the integral in Eq. (19–52) as

$$\begin{aligned}&\left(1 - \frac{c^2}{n^2u^2}\right)\int_{-\infty}^{\infty}\frac{\sin^2\left[\left(1 - \dfrac{nu}{c}\cos\theta\right)\dfrac{\omega X}{u}\right]}{\omega^2\left(1 - \dfrac{nu}{c}\cos\theta\right)^2\Big/ u^2}\, d(\cos\theta) \\ &\qquad = \left(1 - \frac{c^2}{n^2u^2}\right)\frac{cX}{\omega n}\int_{-\infty}^{\infty}\frac{\sin^2 x}{x^2}\, dx = \frac{cX\pi}{\omega n}\left(1 - \frac{c^2}{n^2u^2}\right). \qquad (19\text{–}53)\end{aligned}$$

The radiation loss in the frequency interval $d\omega$ is then

$$\delta W = \frac{e^2 X}{2\pi\epsilon_0 c^2}\left(1 - \frac{c^2}{n^2 u^2}\right)\omega\, d\omega. \tag{19–54}$$

The corresponding energy loss per unit length of path becomes

$$\frac{\delta W}{\delta L} = \frac{e^2}{4\pi\epsilon_0 c^2}\left(1 - \frac{c^2}{n^2 u^2}\right)\omega\, d\omega, \tag{19–55}$$

since the total path considered is $2X$. In terms of the number of quanta of energy $\hbar\omega$, Eq. (19–55) becomes

$$\frac{\delta N}{\delta L} = \frac{e^2}{4\pi\epsilon_0 \hbar c}\left(1 - \frac{c^2}{n^2 u^2}\right)\frac{d\omega}{c} = \alpha\left(1 - \frac{c^2}{n^2 u^2}\right)\frac{d\omega}{c}, \tag{19–56}$$

which is independent of the frequency except insofar as the index of refraction depends on the frequency. Since the velocity of propagation for high frequencies in a dielectric approaches its vacuum value, $n(\omega)$ will become smaller than c/u for sufficiently large values of ω, and the total number of quanta is finite.

SUGGESTED REFERENCES

Radiation from accelerated charges is treated in practically all books on electrodynamics. Three of the best are:

R. Becker, *Theorie der Elektrizität*, Band II.
L. Landau and E. Lifshitz, *The Classical Theory of Fields.*
A. Sommerfeld, *Electrodynamics.*

Briefer accounts will be found in:

J. A. Stratton, *Electromagnetic Theory.* The emphasis here is on antennas and antenna arrays, but the accelerated charge fields are also derived.

W. R. Smythe, *Static and Dynamic Electricity.* The radiation fields of a charge are included in Chapter XIV, together with special relativity.

For a treatment of the Čerenkov effect, see L. I. Schiff, *Quantum Mechanics.*

EXERCISES

1. Equation (19–19) is valid in a proper frame. Show that a Lorentz transformation of *both* sides of the equation leads to Eq. (19–37).

2. In a betatron, electrons are accelerated to velocities comparable to that of light while moving in a circle of fixed radius. Find the energy radiated per one revolution of an electron for $u \sim c$.

3. Show that most of the radiation from a very fast electron is confined to a cone of angle $\theta_0 = W_0/W$, where W is the total energy and $W_0 = m_0 c^2$.

4. Show that the angular distribution of the total radiation emitted when one particle, e_1, m_1, passes another, e_2, m_2 at relative velocity $u \ll c$, is given by

$$-\frac{dW}{d\Omega} = \frac{\text{const}}{b^3}\left(\frac{e_1}{m_1} - \frac{e_2}{m_2}\right)^2 (4 - n_x^2 - 3n_y^2),$$

where n_x and n_y are the components of a unit vector $\mathbf{n}$ directed along the direction of the radiation and b is the impact parameter, assumed constant during the motion. The velocity u is in the x-direction, and the two charges lie in the xy-plane.

5. Consider a charge e passing a heavy charge Ze with impact parameter b. The velocity is not necessarily small. Assuming that the angular change in path is small, calculate the transverse momentum imparted.

6. From the result of problem 5 calculate the mean square scattering angle of a charge passing through a thickness t of material containing N_v charges of magnitude Ze per unit volume. Assume that the impact parameters available range from b_1 to b_2.

7. Calculate the classical cross section (per unit frequency interval of the emitted radiation) for radiation from an electron colliding with a perfectly hard sphere of radius a with $u \ll c$.

8. Estimate the output of Čerenkov radiation for a 350-Mev neutron traveling in glass of refractive index 1.9. (The current associated with the neutron is due to its magnetic dipole moment.) Compare this answer with the Čerenkov radiation from a proton of the same energy under the same conditions.

Hint: Put $\mathbf{j} = \nabla \times \mathbf{M}$ in Eq. (19–48) and integrate by parts.

CHAPTER 20

RADIATION REACTION

It was shown in Chapter 10 that the law of conservation of momentum could be reconciled with an electrical interaction of the form

$$\mathbf{F} = e(\mathbf{E} + \mathbf{u} \times \mathbf{B}) \tag{20-1}$$

only if a momentum of density

$$\mathbf{g} = \frac{\mathbf{N}}{c^2} = \frac{\mathbf{E} \times \mathbf{H}}{c^2} \tag{20-2}$$

is assumed for the electromagnetic field. Similarly, the electromagnetic field carries an energy density

$$U_v = \frac{\mathbf{E} \cdot \mathbf{D} + \mathbf{H} \cdot \mathbf{B}}{2}, \tag{20-3}$$

in order that energy be conserved in a closed system containing both matter and radiation. Let us apply these considerations to the field of an electron.

20–1 Electromagnetic mass. If we consider the energy and momentum balance of the fields due to an electron, we are immediately led to the following qualitative conclusions:

1. If an electron is in uniform motion its field will contribute to the momentum, since for a small virtual change in velocity the momentum of both particle and field would change simultaneously.
2. If an electron radiates by virtue of an acceleration produced by an external force, the external force must supply both the energy and the momentum required by the change in fields. This can be done only by means of a reaction force, produced by action of the radiation field on the electron itself.

To arrive at somewhat more quantitative conclusions, let us first consider the electron in uniform motion with $u \ll c$. If the velocity is changed by a small amount $\delta\mathbf{u}$, the magnetic field changes by

$$\delta\mathbf{B} = \frac{\mu_0}{4\pi} \frac{e\,\delta\mathbf{u} \times \mathbf{r}}{r^3}. \tag{20-4}$$

Consider the magnetic flux passing through an area in a narrow strip normal to $\delta\mathbf{u}$ and extending from the electron to infinity, as indicated in

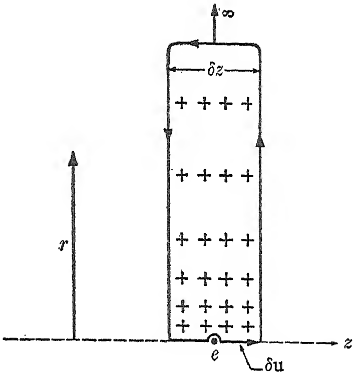


FIG. 20–1 Path of integration for qualitative computation of the reaction of an electron to a small change in velocity.

Fig. 20–1. The flux change of Eq. (20–4) produces an electromotive force around the boundary of the strip. If the z-axis is taken along $\delta\mathbf{u}$, as shown, curl $\mathbf{E} = -\partial\mathbf{B}/\partial t$ reduces to $\partial E_z/\partial r = \partial B/\partial t$, and hence the impulse received by the charge due to the change in velocity $\partial\mathbf{u}$ is approximately

$$\delta g = eE_z\,\delta t = e\int_{\infty}^{r_0} \delta B\,dr = \frac{e^2}{4\pi\epsilon_0 c^2}\frac{\delta u}{r_0}, \tag{20–5}$$

where r_0 is the (unspecified) "electron radius." We see immediately that this impulse would be infinite if the electron were a point charge; on the other hand, if the electron has a finite extent of order r_0, then the magnetic field will react a maximum amount at r_0, and the reaction will go to zero at the center. If we assume a radius r_0, Eq. (20–5) implies that as far as the accelerating force producing $\delta\mathbf{u}$ is concerned there is an effective "mass" of order

$$m \simeq \frac{e^2}{4\pi\epsilon_0 r_0 c^2} \tag{20–6}$$

or, in Gaussian units, $m = e^2/r_0c^2$.

Equation (20–6) was derived by computing the inertial reaction of the field on the driving force. According to the general considerations of Chapter 10, we should obtain the total momentum if we integrate $\mathbf{g} = \mathbf{E} \times \mathbf{H}/c^2$ over all space. Again, if we take the field to be that of a point charge, the integration will diverge, so that a lower limit corresponding to a finite radius r_0 of the electron must be introduced. For low velocities this integration can be carried out very easily, using the quasi-static fields for which

$$\mathbf{H} = \epsilon_0\mathbf{u} \times \mathbf{E}. \tag{20–7}$$

For a symmetrical charge distribution the momentum,

$$\mathbf{G} = \int \frac{\epsilon_0\mathbf{E} \times (\mathbf{u} \times \mathbf{E})}{c^2}\,dv = \epsilon_0\int \frac{\mathbf{u}E^2 - (\mathbf{u}\cdot\mathbf{E})\mathbf{E}}{c^2}\,dv, \tag{20–8}$$

is axially symmetric about the direction of the velocity, and it is clear that $\mathbf{G}$ itself is along $\mathbf{u}$, which we may take in the x-direction. The second term in the integrand is therefore just $\mathbf{u}E_x^2$. For low velocities, $\mathbf{E}$ is spherically symmetric, and

$$\mathbf{G} = \int_{r_0}^{\infty} \frac{\epsilon_0(E_y^2 + E_z^2)\mathbf{u}\,dv}{c^2} = \frac{2\mathbf{u}}{3}\int_{r_0}^{\infty} \frac{\epsilon_0 E^2\,dv}{c^2} = \frac{4}{3}\frac{U_0}{c^2}\mathbf{u}, \tag{20–9}$$

where

$$U_0 = \tfrac{1}{2}\int \mathbf{E}\cdot\mathbf{D}\,dv \tag{20–10}$$

is the electrostatic field energy of the charge e. If, for instance, it is assumed that the charge is distributed uniformly over a spherical shell of radius r_0,

$$U_0 = \frac{1}{2}\int_{r_0}^{\infty} \frac{\epsilon_0 e^2 4\pi r^2\,dr}{(4\pi\epsilon_0 r^2)^2} = \frac{e^2}{8\pi\epsilon_0 r_0}, \tag{20–11}$$

and thus Eq. (20–9) is in qualitative agreement with an electromagnetic mass as given by Eq. (20–6).

At high velocities the fields of Chapter 18, with $s \neq r$, must be used in the computation of $\mathbf{G}$. The resultant relation between momentum and velocity depends on the behavior of the shape of the electron during its motion. If it is assumed to suffer a Lorentz contraction in the direction of its motion (an assumption due to Lorentz, and later made plausible by relativistic arguments), the integration is relatively straightforward, and the final result is

$$\mathbf{G} = \frac{4}{3}\frac{U_0}{c^2}\frac{\mathbf{u}}{\sqrt{1 - u^2/c^2}}, \tag{20–12}$$

where U_0 is given by Eq. (20–10).

With the exception of the factor $\frac{4}{3}$, Eq. (20–12) is consistent with the general relations

$$\mathbf{G} = m\mathbf{u},$$

$$m = \frac{m_0}{\sqrt{1 - u^2/c^2}}, \tag{20–13}$$

$$U_0 = m_0 c^2,$$

which we have found to be the correct relativistic expressions relating mass, energy, and momentum. The factor $\frac{4}{3}$ is significant, however, and indicates that the purely "electromagnetic mass" calculated here is not the actual mass, although the ratio of electromagnetic mass to total field energy is independent of r_0 and hence independent of any approximations made in the fields. An additional mass $-U_0/3c^2$, whose origin is not electromagnetic, is needed to account for the observed quantities, which do obey the relativistically correct equations (20–13). This extra mass (or energy) presumably represents the nonelectromagnetic binding which must be present to make the charge system of the electron stable.

20–2 The reaction necessary to conserve radiated energy. We have seen that physically the mass associated with the momentum of the field is due to a reaction on whatever tends to change the velocity of an electron in steady motion. Let us now obtain the reaction force which must be present if the electron is accelerated, giving rise to a rate of energy loss

$$-\frac{dW}{dt} = \frac{e^2\dot{u}^2}{6\pi\epsilon_0 c^3}. \tag{19–19}$$

No essential restriction is implied by the form of Eq. (19–19), since the results can be generalized by the appropriate transformation of Chapter 17 to be valid for all allowable velocities. For energy to be conserved, the reaction force should satisfy the condition

$$\mathbf{F}\cdot\mathbf{u} + \frac{e^2}{6\pi\epsilon_0}\frac{\dot{u}^2}{c^3} = 0. \tag{20–14}$$

This equation obviously has no solution for **F** which can be instantaneously correct for all times, since **u** and $\dot{\mathbf{u}}$ are basically uncorrelated. We must be content with a solution representing an average over a sufficiently long period of time. This means that we have only an average energy balance between the force and the radiation field, and that an extra fluctuation will be available which will be stored in the induction field. On this basis, Eq. (20–14) becomes

$$\int_{t_1}^{t_2}(\mathbf{F}\cdot\mathbf{u})\,dt + \int_{t_1}^{t_2}\frac{e^2}{6\pi\epsilon_0 c^3}\,\dot{u}^2\,dt = 0. \tag{20–15}$$

Integrating by parts, we obtain

$$\int_{t_1}^{t_2}\left(\mathbf{F} - \frac{e^2\ddot{\mathbf{u}}}{6\pi\epsilon_0 c^3}\right)\cdot\mathbf{u}\,dt + \left[\frac{e^2\mathbf{u}\cdot\dot{\mathbf{u}}}{6\pi\epsilon_0 c^3}\right]_{t_1}^{t_2} = 0. \tag{20–16}$$

The integrated term represents the "fluctuation" referred to above; for periodic motion, or for an acceleration occurring over a limited time, it will not affect the integrated energy balance. On the average, energy will be conserved if we put

$$\mathbf{F}_{\text{rad}} = \frac{e^2\ddot{\mathbf{u}}}{6\pi\epsilon_0 c^3} \tag{20–17}$$

as the radiation reaction force.

The reaction is in addition to that demanded by the conservation of momentum. The total reaction on the electron itself would then be

$$\mathbf{F} = \frac{e^2\ddot{\mathbf{u}}}{6\pi\epsilon_0 c^3} - m_{\text{el. mag.}}\,\dot{\mathbf{u}}, \tag{20–18}$$

where $m_{\text{el. mag.}}$ is the electromagnetic mass

$$m_{\text{el. mag.}} = \frac{4}{3}\frac{U_0}{c^2}. \tag{20–19}$$

20–3 Computation of radiation reaction from the fields. It is instructive to compute Eq. (20–18) by direct integration of the interaction of the radiation field of one part of the electron with the other parts. This calculation, originally due to Lorentz, shows very clearly the limitations of the theory.

For this calculation we shall make the following assumptions:

1. We shall choose a frame such that the element de of the electron, on which another element de' acts, is at rest: $\mathbf{u}(t)_{de} = 0$.

2. None of the quantities $\mathbf{u}$, $\dot{\mathbf{u}}$, $\ddot{\mathbf{u}}$, etc., changes very much during the time it takes for an electromagnetic signal to cross the electron. This is equivalent to the conditions that $u \ll c$, $\dot{u} \ll c^2/r_0$, $\ddot{u} \ll \dot{u}c/r_0$, etc. The solution is effectively a power series in $\tau_0 = r_0/c = e^2/4\pi\epsilon_0 mc^3$ times the operator $\partial/\partial t$.

3. The fields will be derived from the Liénard-Wiechert potentials, even though their correctness at distances of the order of r_0 is doubtful. Therefore, only terms not containing r_0 explicitly will be considered as having physical significance.

4. The electron charge distribution has spherical symmetry.

Consider an element of charge de at x_α affected by the element de' at x'_α, as shown in Fig. 20–2. From Eq. (19–13), we have the electric field at x due to de':

$$4\pi\epsilon_0\, d\mathbf{E}(t) = \frac{de'}{s^3}\left\{\frac{1}{c^2}\mathbf{r}\times\left[\left(\mathbf{r} - \frac{\mathbf{u}(t')r}{c}\right)\times\dot{\mathbf{u}}(t')\right] + \left[1 - \left(\frac{u(t')}{c}\right)^2\right]\left(\mathbf{r} - \frac{\mathbf{u}(t')r}{c}\right)\right\}. \tag{20–20}$$

Since the field is expressed in terms of the electron's condition at t', the problem would be insoluble without knowledge of the electron's entire past if we allow arbitrary motions. Treatment of the problem is made possible only by the restrictions of item 2 above. With these limitations on the motion, it is possible to refer all velocities and accelerations to the time of arrival of the signal at x_α, which must be common to all elements de'. Only in this way are we able to integrate over the entire electron. The change from t' to t is accomplished by expanding the functions of $t' = t - r/c$ in powers of r/c, remembering that $\mathbf{u}(t) = 0$, in accord with assumption 1.

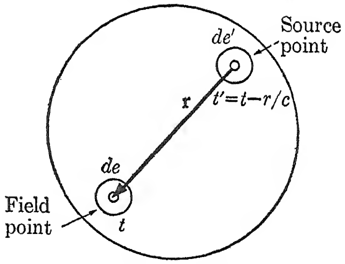


FIG. 20-2 For computing the interaction between one part of the electron and the other parts.

$$\dot{\mathbf{u}}(t') = \dot{\mathbf{u}}(t) - \frac{r}{c}\ddot{\mathbf{u}}(t) + \cdots, \tag{20-21}$$

$$\mathbf{u}(t') = -\frac{\dot{\mathbf{u}}(t)r}{c} + \frac{\ddot{\mathbf{u}}(t)}{2}\frac{r}{c} - \cdots, \tag{20-22}$$

$$s^{-3} = r^{-3}\left[1 - \frac{3\dot{\mathbf{u}}(t)\cdot\mathbf{r}}{c^2} + \frac{3r}{2}\left(\frac{\ddot{\mathbf{u}}(t)\cdot\mathbf{r}}{c^3}\right)\cdots\right]. \tag{20-23}$$

Let us carry terms only to order $(r/c)^3$; thus $[u(t')/c]^2$ can be neglected, and the resulting expression is considerably simplified. In terms of t, Eq. (20-20) then becomes

$$\begin{aligned} 4\pi\epsilon_0\, d\mathbf{E} &= de'\left[\frac{\mathbf{r}(\mathbf{r}\cdot\dot{\mathbf{u}})}{r^3c^2} - \frac{\dot{\mathbf{u}}}{rc^2} - \frac{\mathbf{r}(\mathbf{r}\cdot\ddot{\mathbf{u}})}{r^2c^3} + \frac{\ddot{\mathbf{u}}}{c^3} + \frac{\mathbf{r}}{r^3}\right. \\ &\qquad \left. - \frac{3\mathbf{r}(\dot{\mathbf{u}}\cdot\mathbf{r})}{r^3c^2} + \frac{3}{2}\frac{\mathbf{r}}{r^2c^3}(\ddot{\mathbf{u}}\cdot\mathbf{r}) + \frac{\dot{\mathbf{u}}}{rc^2} - \frac{\ddot{\mathbf{u}}}{2c^3}\right] \\ &= de'\left[-\frac{2\mathbf{r}(\dot{\mathbf{u}}\cdot\mathbf{r})}{r^3c^2} + \frac{1}{2}\frac{\mathbf{r}(\ddot{\mathbf{u}}\cdot\mathbf{r})}{r^2c^3} + \frac{\mathbf{r}}{r^3} + \frac{\ddot{\mathbf{u}}}{2c^3}\right]. \end{aligned} \tag{20-24}$$

It will be easier to take advantage of our assumption of spherical symmetry if we write Eq. (20-24) in tensor notation:

$$dE_\alpha = \frac{de'}{4\pi\epsilon_0}\left[-\frac{2r_\alpha(\dot{u}_\beta r_\beta)}{r^3c^2} + \frac{1}{2}\frac{r_\alpha(\ddot{u}_\beta r_\beta)}{r^2c^3} + \frac{r_\alpha}{r^3} + \frac{\ddot{u}_\alpha}{2c^3}\right]. \tag{20-25}$$

Now the average of r_α over a spherically symmetric region is zero, and the average of $r_\alpha r_\beta$ is simply $\frac{1}{3}r^2\,\delta_{\alpha\beta}$. Therefore the third term of Eq. (20-25) will contribute nothing to the final integral, and we may write

$$\overline{dE_\alpha} = \frac{de'}{4\pi\epsilon_0}\left(-\frac{2}{3}\frac{\dot{u}_\alpha}{c^2r} + \frac{2}{3}\frac{\ddot{u}_\alpha}{c^3}\right). \tag{20-26}$$

Integrating over de' and de, we obtain for the reaction force

$$\mathbf{F} = \iint de\, d\mathbf{E} = \frac{e^2\ddot{\mathbf{u}}}{6\pi\epsilon_0 c^3} - m_{\text{el. mag.}}\dot{\mathbf{u}}, \tag{20-27}$$

where

$$m_{\text{el. mag.}} = \frac{4}{3c^2}\iint \frac{de\, de'}{8\pi\epsilon_0 r} = \frac{4}{3c^2}U_0, \tag{20-28}$$

with U_0 again representing the electrostatic energy of the electron in its own field. This is in complete agreement with Eq. (20–18), the result required by the conservation laws.

In Eq. (20–27) the term proportional to $\ddot{\mathbf{u}}$ is independent of the extent of the electron, and therefore presumably independent of the detailed structure of the electron. The mass term is indeed structure-dependent, but its relation to the electrostatic energy is not. If we had carried the calculation to higher order in r/c the additional terms would form an ascending series in r_0 which, according to our assumption 3, could not be assumed to have physical significance.

20–4 Difficulties in classical electron theory. The direct computation of the reaction force indicates certain unsolved problems in the classical theory of the electron:

1. The theory gives a relation between electrostatic field energy and mass, but does not indicate that either of them should be finite for a point field source.

2. The electromagnetic mass of the electron does not account for the entire mass; the electron must have nonelectromagnetic mass of unknown origin to account for its stability.

3. The restriction $\ddot{u} \ll \dot{u}c/r_0$, which was necessary to permit the calculation of Eq. (20–27), is equivalent to putting

$$|m\dot{\mathbf{u}}| \gg \left|\frac{e^2\ddot{\mathbf{u}}}{6\pi\epsilon_0 c^2 r_0}\frac{r_0}{c}\right| = |\mathbf{F}_{\text{rad}}|. \tag{20-29}$$

This means that there must be an external force large compared with the radiation reaction force in order for the theory to be valid. Hence the force equations derived here apply only to electrons whose behavior is governed by external forces. If we attempt to consider the force of radiation on a free particle, the equation of motion takes the form

$$k\ddot{\mathbf{u}} - \dot{\mathbf{u}} = 0, \tag{20-30}$$

where

$$k \simeq \tau_0 = \frac{e^2}{4\pi\epsilon_0 mc^3}.$$

The solution of Eq. (20–30) is

$$u = u_0 e^{t/k} + \text{constant} \times t. \tag{20–31}$$

Now τ_0 is roughly 10^{-23} second for an electron, so that if Eq. (20–31) were justified the particle would go shooting off to infinity.

Quantum mechanics has not solved these difficulties. In fact, problems arise concerning the charge that are similar to the classical difficulties in regard to electronic mass. The experimental values of e and m can, of course, be used to describe the external behavior of the electron; the unsolved problems are those concerned with the details of electronic structure.

SUGGESTED REFERENCES

H. A. Lorentz, *The Theory of Electrons.* The proofs of the radiation reaction formula are to be found in the notes at the end.

L. Landau and E. Lifshitz, *The Classical Theory of Fields.* Chapter 9 includes radiation damping.

A. Sommerfeld, *Electrodynamics.* Section 36 is devoted to an excellent discussion of radiation reaction.

Discussions of electromagnetic mass are to be found in:

R. Becker, *Theorie der Elektrizität*, Band II.

J. H. Jeans, *The Mathematical Theory of Electricity and Magnetism.*

M. Mason and W. Weaver, *The Electromagnetic Field.*

EXERCISES

1. Find the radiation reaction force on an electron being accelerated in a betatron (i.e., while moving in a circle of fixed radius). What is the dependence of the reaction force on the particle energy?

2. Given an electron moving initially in a circle of radius a about a proton. If energy were lost by radiation according to the classical law, what would happen? Find how long the system would last.

3. Two parallel dipole antennas are excited sinusoidally 90° out of phase and separated by a quarter wavelength in space. Find the net force on the antennas.

CHAPTER 21

RADIATION, SCATTERING, AND DISPERSION

21–1 Radiative damping of a charged harmonic oscillator. Provided that none of the restrictions of Section 20–3 is violated, the theory of radiation reaction can be applied with some confidence to a bound electron. Consider an electron bound with a harmonic force $\mathbf{F} = -k\mathbf{x}$, corresponding to a natural frequency $\omega_0 = \sqrt{k/m}$. The equation of motion, including the self-force of the electron, is then

$$\ddot{x} + \omega_0^2 x = \frac{e^2 \dddot{x}}{6\pi\epsilon_0 c^3 m} = \frac{2}{3}\tau_0 \dddot{x}. \tag{21–1}$$

(The mass reaction term is implicit here, i.e., m is the empirically determined mass, including the electromagnetic mass.) As we have seen in Eq. (20–29), the right side of Eq. (21–1) must be small in comparison with the binding term. In that case, we can make the approximation

$$\dddot{x} \simeq -\omega_0^2 \dot{x}. \tag{21–2}$$

If for convenience we set

$$\gamma = \frac{2}{3}\tau_0\omega_0^2 = \frac{e^2\omega_0^2}{6\pi\epsilon_0 c^3 m}, \tag{21–3}$$

then Eq. (21–1) becomes

$$\ddot{x} + \gamma\dot{x} + \omega_0^2 x = 0. \tag{21–4}$$

The solution of Eq. (21–4), valid for small γ, is

$$x = Ae^{-i\omega_0 t}e^{-\gamma t/2}. \tag{21–5}$$

The energy of the oscillator thus falls off as $e^{-\gamma t}$,

$$W \simeq \frac{mA^2\omega_0^2}{2}\, e^{-\gamma t} \tag{21–6}$$

and

$$-\frac{dW}{dt} = \gamma W = \frac{e^2\omega_0^4 A^2}{12\pi\epsilon_0 c^3}\, e^{-\gamma t}. \tag{21–7}$$

But Eq. (21–7) is just the rate of radiation, averaged over a cycle, so that energy is conserved in accord with our first derivation of $\mathbf{F}_{\text{rad}}$. Equation (21–7) corresponds to a damped wave train emitted by the oscillator after a given amplitude A has been excited by an external impulse. It is clear from Eq. (21–7) that $1/\gamma$ is the mean duration of the radiated pulse when it is averaged over the energy. This is the classical quantity corresponding to the quantum mechanical "lifetime" of an excited state produced by an external impulse.

The limitation on the frequency imposed by the condition that $\gamma \ll \omega_0$ is entirely unimportant except for high-energy γ-rays. In terms of the quantum energy of the outgoing radiation, the inequality becomes

$$\hbar\omega_0 \ll \frac{4\pi\epsilon_0\hbar c}{e^2} mc^2 = 137\, mc^2 \simeq 70 \text{ Mev.} \tag{21-8}$$

Because of the radiative damping, the radiation emitted by the oscillator is not monochromatic. The line width can be obtained by making a Fourier analysis of the fields, which are proportional to Eq. (21-5). If

$$E = \int_{-\infty}^{\infty} E_\omega e^{-i\omega t}\, d\omega = E_0 e^{-i\omega_0 t} e^{-\gamma t/2}, \tag{21-9}$$

then

$$E_\omega = \frac{E_0}{2\pi} \int_0^\infty e^{-i\omega_0 t} e^{-\gamma t/2} e^{i\omega t}\, dt$$

$$= \frac{E_0}{2\pi} \frac{1}{i(\omega - \omega_0) - \gamma/2}. \tag{21-10}$$

Equation (21-10) corresponds (see Fig. 21-1) to a radiation intensity

$$I_\omega = \frac{I_0\gamma}{2\pi} \frac{1}{(\omega - \omega_0)^2 + \gamma^2/4} \tag{21-11}$$

normalized in such a way that

$$\int_{-\infty}^{\infty} I_\omega\, d\omega = I_0. \tag{21-12}$$

The full frequency width at half intensity is therefore

$$\Delta\omega \simeq \gamma = \tfrac{2}{3}\tau_0\omega_0^2. \tag{21-13}$$

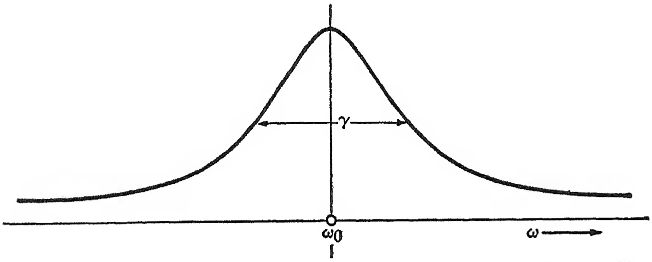


FIG. 21-1 Radiation intensity plotted against frequency as given in Eq. (21-11)

The corresponding width expressed in wavelength is

$$\Delta\lambda = \frac{2\pi\,\Delta\omega c}{\omega_0^2} = \frac{4\pi}{3}c\tau_0 = \frac{4\pi}{3}r_0 \simeq 10^{-12}\text{ cm}, \tag{21-14}$$

which is constant, independent of the frequency of the oscillator. In practice, the radiation damping is, of course, not the only source of width for a spectral line: interruptions of the wave train by atomic collisions and Doppler shifts also contribute to line width.

The relation between line width and mean life given by Eq. (21–13), namely, $\Delta\omega\gamma^{-1} = 1$, is equivalent to the relation

$$\Delta E\,\Delta t \sim \hbar, \tag{21-15}$$

where $\Delta E = \hbar\,\Delta\omega$. Equation (21–15) is the quantum mechanical relation between the lifetime and the energy width of a state.

21–2 Forced vibrations. In the previous section, we have discussed the motion and consequent radiation of a bound electron following a transient disturbing impulse; this is the classical theory of spectral emission. Let us now consider the steady state motion of a harmonically bound electron forced to vibrate by an external electromagnetic wave. The classical theories of the scattering, absorption, and dispersion of light are based on such a consideration.

For the treatment of optical frequencies and even soft x-rays, we may assume that the velocity of the electron is well below the velocity of light, and therefore we may neglect the effect of the magnetic vector in the incoming radiation. The equation of motion of a bound electron in an external (polarized) plane wave field $\mathbf{E} = \mathbf{E}_0 e^{-i\omega t}$ is then

$$\ddot{\mathbf{x}} - \frac{2}{3}\tau_0\dddot{\mathbf{x}} + \omega_0^2\mathbf{x} = \frac{e}{m}\mathbf{E}_0 e^{-i\omega t}, \tag{21-16}$$

with τ_0 as defined in Eq. (21–1). We shall be interested in the steady state solution, which is proportional to $e^{-i\omega t}$, and so we may put

$$\dddot{\mathbf{x}} \simeq -\omega^2\dot{\mathbf{x}}. \tag{21-17}$$

We may define

$$\gamma = \tfrac{2}{3}\tau_0\omega^2, \tag{21-18}$$

although it does not matter very much whether we use Eq. (21–18) or Eq. (21–3); τ_0 is sufficiently small so that the term is important only when $\omega \sim \omega_0$, i.e., only near resonance. The steady state solution of Eq. (21–16) is then given by

$$\mathbf{x} = \mathbf{E}_0\frac{e}{m}\left(\frac{1}{\omega_0^2 - \omega^2 - i\omega\gamma}\right)e^{-i\omega t}. \tag{21-19}$$

The acceleration of the charge in terms of the external field,

$$\ddot{\mathbf{x}} = \frac{e}{m}\left(\frac{-\omega^2}{\omega_0^2 - \omega^2 - i\omega\gamma}\right)\mathbf{E}, \tag{21-20}$$

may be substituted in Eqs. (19–20) for the dipole radiation fields to give the scattered radiation. We shall examine in detail several applications of Eq. (21–20).

21–3 Scattering by an individual free electron.

For an unbound and weakly accelerated electron, $\gamma \simeq 0$, $\omega_0 \simeq 0$, so that

$$\ddot{\mathbf{x}} = \frac{e}{m}\mathbf{E}. \tag{21-21}$$

This acceleration gives rise to a radiation field of magnitude

$$E(\mathbf{r}) = \frac{e \sin \alpha \ddot{x}}{4\pi\epsilon_0 r c^2}, \tag{21-22}$$

where α is the angle between $\mathbf{E}$ and $\mathbf{r}$, as shown in Fig. 21–2. The rate at which the re-radiated energy crosses a unit area is

$$|\mathbf{N}| = \sqrt{\frac{\epsilon_0}{\mu_0}}\left[\frac{e \sin \alpha \ddot{x}}{4\pi\epsilon_0 r c^2}\right]^2, \tag{21-23}$$

$$|\mathbf{N}| = r_0^2 \sin^2 \alpha \left(\frac{I_0}{r^2}\right), \tag{21-24}$$

where $r_0 = e^2/4\pi\epsilon_0 mc^2$ is the classical electron radius, and

$$I_0 = \sqrt{\frac{\epsilon_0}{\mu_0}}\, E^2 \tag{21-25}$$

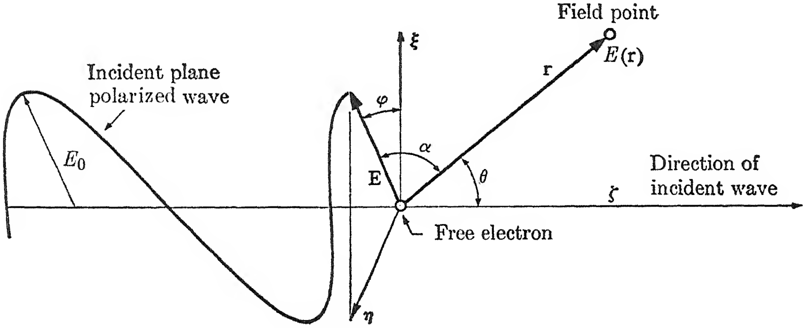


FIG. 21–2 A plane wave to be scattered by an electron, showing the polarization and scattering angles.

is the primary intensity. Then

$$-\frac{dW}{dt} = \int r^2 N\, d\Omega = \frac{8\pi I_0 r_0^2}{3}. \tag{21–26}$$

If we divide Eq. (21–26) by I_0, the intensity of the incoming radiation, we obtain the effective scattering cross section per electron,

$$\sigma_0 = \frac{8\pi}{3} r_0^2. \tag{21–27}$$

Equation (21–26) is known as the Thomson scattering formula, and Eq. (21–27) is called the cross section for Thomson scattering. Both are constant, independent of the frequency.

Comparison with experiment is facilitated if we express Eq. (21–24) in terms of the scattering angle θ and the polarization angle φ, shown in Fig. 21–2. The angles α, θ, and φ are related by

$$\cos\alpha = \cos\varphi \sin\theta,$$

i.e., $\sin^2\alpha = 1 - \cos^2\varphi(1 - \cos^2\theta)$.

If the primary wave is unpolarized, i.e., randomly polarized, we must average over φ, and thus we obtain

$$\overline{\sin^2\alpha} = \tfrac{1}{2}(1 + \cos^2\theta),$$

since the average of $\cos^2\varphi$ is $\frac{1}{2}$. In terms of θ, the total rate of scattered energy is

$$-\frac{dW}{dt} = I_0 r_0^2 \int \frac{(1 + \cos^2\theta)}{2}\, d\Omega = \frac{8\pi I_0 r_0^2}{3}. \tag{21–28}$$

The differential scattering cross section per unit solid angle is given by

$$\frac{d\sigma_0}{d\Omega} = \frac{r_0^2(1 + \cos^2\theta)}{2}. \tag{21–29}$$

This cross section is shown graphically as a function of the scattering angle in Fig. 21–3. The quantum mechanical formula for the scattered radiation approaches Eq. (21–29) as the frequency goes to zero, but the scattered intensity is not symmetrical. In general, the scattered radiation is more concentrated in the forward direction, as indicated by the curve for radiation whose primary energy is 100 kilo-electron-volts.

There is another feature of Thomson scattering which is modified by quantum considerations. Classically, the scattered radiation has the same frequency as that of the incoming wave. One qualification must be made: momentum, as well as energy, is removed from the incident beam, and from the symmetry of the scattered radiation it is evident that this momentum is given to the electron. This is equivalent to an average force in

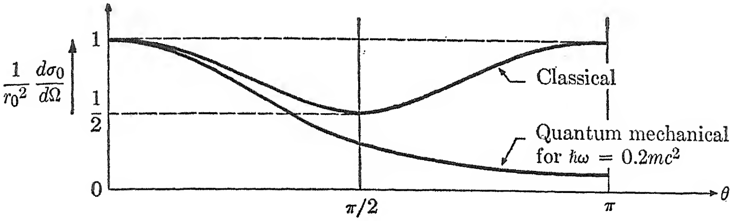


FIG. 21-3 Differential scattering cross section as a function of scattering angle, showing deviations from Thomson scattering for moderately high frequencies.

the direction of the incoming beam, and equal to $(I_0/c)\sigma_0$, which would gradually impart a net velocity to the electron. Actually this is not observed. What happens instead is that each quantum of radiation imparts some of its energy and momentum to the electron, and that the frequency of the scattered quantum depends on the angle of scattering. For a discussion of Compton scattering and the quantum mechanical Klein-Nishina formula, see, for example, W. Heitler, *The Quantum Theory of Radiation*, second edition, pp. 149–161.

21–4 Scattering by a bound electron. For the general case of a harmonically bound electron, the acceleration given in Eq. (21–20) must be used in Eq. (21–23) to determine the re-radiated energy per unit area in a given direction. With this substitution, the total cross section for scattering becomes

$$\sigma = \sigma_0 \frac{\omega^4}{(\omega_0^2 - \omega^2)^2 + (\gamma\omega)^2}, \tag{21–30}$$

where σ_0 is the Thomson cross section given in Eq. (21–27). The angular distribution is the same as in the free electron case.

The maximum value of Eq. (21–30) is obtained if $\omega \simeq \omega_0$, i.e., for resonance scattering. In that case, the scattering cross section can become very large: the cross section

$$\sigma = \sigma_0 \left(\frac{\omega_0}{\gamma}\right)^2 \simeq 6\pi\lambda\!\!\!^{-}{}_0^2 \tag{21–31}$$

is greatly in excess of the classical area of the electron.

For strong binding, $\omega \ll \omega_0$, $\gamma \ll \omega_0$, and Eq. (21–30) becomes

$$\sigma = \sigma_0 \left(\frac{\omega}{\omega_0}\right)^4, \tag{21–32}$$

giving a cross section that depends on the inverse fourth power of the incident wavelength. In Chapter 12 we obtained scattering from a conducting sphere which behaved in the same way for the case that the dimensions of the sphere were small compared with λ, and it is found that a small dielectric sphere scatters radiation according to the same frequency dependence. The inverse fourth power law was first derived by Rayleigh in his investigation of the blueness of the sky; he assumed that the individual molecules scatter in a wholly random fashion, so that the phases are random and the intensities add. Not only does Eq. (21–32) account for the blue color; with average empirical values of ω_0 it leads quantitatively to generally good agreement with observed atmospheric scattering, both in total amount and in polarization. Justification for the assumption of random phases will be discussed later.

21–5 Effect of a volume distribution of scatterers. If we have N electronic scatterers per unit volume, each of them will scatter an incoming plane wave in accord with Eq. (21–30). The scattered radiation will, in general, combine coherently with the external field, and thus modify the effective velocity of the wave. It is possible to obtain the effect of a great number of scatterers, electrons or molecules, by superposition of the scattered wavelets. For the study of refraction, however, it is mathematically simpler to treat the re-radiation as being due to the electric polarization of an entire volume element, rather than to the combination of single electron displacements. The polarization is related to the individual electron displacement by

$$\mathbf{P} = Ne\mathbf{x}. \tag{21–33}$$

The effect of the volume polarization is to add the polarization current $\partial\mathbf{P}/\partial t$ to the vacuum displacement current $\epsilon_0\,\partial\mathbf{E}/\partial t$ as a circulation source of magnetic field. Equation (9–6)(4) becomes

$$\nabla \times \mathbf{B} = \mu_0\left(\epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \frac{\partial \mathbf{P}}{\partial t}\right). \tag{21–34}$$

This equation, when combined with $\nabla \times \mathbf{E} = -\partial\mathbf{B}/\partial t$ in the usual way, yields the homogeneous wave equations

$$\begin{aligned} \nabla^2\mathbf{E} - \frac{n^2}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2} &= 0, \\ \nabla^2\mathbf{B} - \frac{n^2}{c^2}\frac{\partial^2\mathbf{B}}{\partial t^2} &= 0, \end{aligned} \tag{21–35}$$

where

$$n = \sqrt{1 + \frac{|\mathbf{P}|}{\epsilon_0|\mathbf{E}|}} = \sqrt{\kappa} \tag{21–36}$$

is the index of refraction.

For dilute systems, in which the scattered energy is very small compared with that in the incident beam, it is a good approximation to take as the local field at the scatterer just the field of the incoming wave. Therefore, from Eq. (21-19),

$$n^2 = \kappa = 1 + \frac{Ne^2}{\epsilon_0 m} \frac{1}{\omega_0^2 - \omega^2 - i\omega\gamma}. \tag{21-37}$$

Equation (21-37) is the dispersion formula for a dilute gas of electrons with a single binding frequency. If $\omega < \omega_0$, we have what is called "normal dispersion," in which the refractive index diminishes with increasing wavelength. At $\omega \simeq \omega_0$ there is an absorption band, at which n has an appreciable imaginary part. On the short wavelength side of a resonance frequency $n < 1$, but we note that c/n is the *phase* velocity: the group velocity, $d\omega/dk$, is less than c on both sides of the resonance frequency. For frequencies in the immediate neighborhood of ω_0 the question of transmission is more complicated, and we shall consider it further in Section 21-6.

If the medium is not dilute, the local field is not equal to the external field, but may be approximated by Eq. (2-36):

$$\mathbf{E}_{\text{eff}} = \mathbf{E} + \mathbf{P}/3\epsilon_0.$$

It is this field that must be substituted in Eq. (21-9) to obtain the polarization as defined in Eq. (21-33). When the fields are eliminated, we obtain the analog of the Clausius-Mossotti relation, Eq. (2-39):

$$\frac{\kappa - 1}{\kappa + 2} = \frac{n^2 - 1}{n^2 + 2} = \frac{Ne^2}{3\epsilon_0 m} \frac{1}{\omega_0^2 - \omega^2 - i\omega\gamma}. \tag{21-38}$$

The resemblance to Eq. (2-39) is even closer if we express Eq. (21-38) in terms of the classical electron radius, r_0, and the wave numbers $k_0 = \omega_0/c$, $k = \omega/c$, $\Gamma = \gamma/c$, and put $N = gN_0/M$ in terms of the density g, the molecular weight M, and Avogadro's number N_0:

$$\frac{3M}{g}\left(\frac{n^2 - 1}{n^2 + 2}\right) = N_0\alpha = \frac{4\pi N_0 r_0}{k_0^2 - k^2 - ik\Gamma}. \tag{21-39}$$

For a given frequency, the right side of Eq. (21-39) depends only on the atomic properties of the material and is thus independent of the density. Since it is referred to a gram molecule, the quantity on the left is called the "molar refraction," and we see that it is equal to the polarizability per gram molecule.

Equations (21-37), (21-38), and (21-39) can be generalized to take account of the fact that all electrons may not have the same binding energy.

If a fraction of the electrons f_i has a binding frequency ω_{0i} and a damping width γ_i, Eq. (21–38), for example, becomes

$$\frac{n^2-1}{n^2+2} = \frac{Ne^2}{3\epsilon_0 m}\sum_i \frac{fi}{\omega_{0i}^2-\omega^2+i\omega\gamma_i}. \tag{21–40}$$

There are several practical examples of the case where the volume distribution consists of free electrons: refraction of electromagnetic waves by the ionosphere, re-radiation by a "plasma" in an electric discharge, and even the refraction of x-rays, since the resonance frequencies of electrons in light atoms are so small in comparison with x-ray frequencies that the electrons may be considered free. For free electron systems that are dilute, Eq. (21–37) becomes

$$n^2 = 1 - \frac{Ne^2}{\epsilon_0 m\omega^2}. \tag{21–41}$$

In terms of the reduced wavelength $\lambda\!\!\!^- = c/\omega$, and the classical electron radius $r_0 = e^2/4\pi\epsilon_0 mc^2$, Eq. (21–40) may be written

$$n^2 = 1 - 4\pi N r_0 \lambda\!\!\!^-{}^2. \tag{21–42}$$

The index of refraction of a volume distribution of free electrons is always less than unity, although the group velocity is less than c. However, for wavelengths longer than a certain value, $\lambda\!\!\!^- > 1/\sqrt{4\pi N r_0}$, n^2 is negative and the refractive index pure imaginary. For such frequencies free electron media can exhibit total reflection at all angles, behaving somewhat like a wave guide for frequencies lower than the cutoff frequency.

21–6 Velocities of propagation. The velocity that appears in the wave equation, such that the plane wave solution is $f(x \pm vt)$, is the phase velocity of the wave, and the index of refraction is defined by $n = c/v$. If all frequencies are propagated at the same velocity there is no dispersion, and c/n is the only velocity involved. If ω_0 is much greater than the radiation frequency, Eqs. (21–37) and (21–38) very nearly fulfill this requirement. Ordinarily, however, since actual radiation consists of more than a single frequency, its passage through a medium is modified by the differences in phase velocity among its components.

In Section 12–6 we arrived at a definition of group velocity by superposing two waves of slightly different frequencies for which the phase velocities are different. The concept of group velocity also applies to a pulse of waves, or a "wave packet," with a continuous spectrum confined to a narrow band of frequencies. Let $\psi(x,t)$ stand for **E** or **B** corresponding to such a narrow group, which can be represented by a Fourier integral

$$\psi = \int a(k)e^{i(kx-\omega t)}\,dk. \tag{21–43}$$

If the amplitude $a(k)$ differs from zero only in the neighborhood of a particular frequency designated by ω_1, k_1, we may write

$$\psi = A(x,t)e^{i(k_1x-\omega_1t)}, \tag{21-44}$$

where

$$A = \int a(k)e^{i[(\Delta k)x-(\Delta\omega)t]}\,dk. \tag{21-45}$$

If A is slowly varying, the group may be considered as a wave of angular frequency ω_1 propagated with amplitude A. The velocity with which A travels corresponds to having the amplitude be constant to an observer moving with that velocity. Hence,

$$\frac{dA}{dt} = \frac{\partial A}{\partial t} + \frac{\partial A}{\partial x}\left(\frac{dx}{dt}\right) = 0,$$

from which

$$\left(\frac{dx}{dt}\right)_{A\text{ constant}} = -\frac{\partial A/\partial t}{\partial A/\partial x} = \frac{\Delta\omega}{\Delta k} \rightarrow \left(\frac{d\omega}{dk}\right)_{\omega=\omega_1}, \tag{21-46}$$

in agreement with Eq. (12–35) for the group velocity, v_g. The expression for v_g may be written in a variety of useful ways:

$$v_g = \frac{d\omega}{dk} = v - \lambda\frac{dv}{d\lambda} = \frac{1}{dk/d\omega}. \tag{21-47}$$

If the phase velocity is a slowly varying function of the frequency, a pulse may travel through a dispersive medium with relatively little change, and the group velocity, defined for its average of maximum amplitude frequency, is the rate of propagation of energy. In the region of anomalous dispersion, however, $dv/d\lambda$ is negative, and the group velocity may easily be greater than c. For any particular frequency it is more reasonable to define the mean velocity of energy transport as the ratio of the time average of the Poynting vector to that of the energy density. This velocity, which is called the energy velocity, actually has a minimum at ω_0.

The practical difficulty of describing the transmission of a signal in the region of anomalous dispersion arises from the fact that the group, or packet, is highly distorted, so that the concept of group velocity as defined above is no longer valid. The problem corresponding to a plane wave that starts through a medium at a particular time has been worked out by Sommerfeld and by Brillouin, and details may be found in Volume II of *Congrès International d'Electricité*, Paris, 1932.* The definition of a signal velocity depends on both the nature of the pulse and the nature of the signal detector, but careful analysis shows that it is never greater than c. Figure 21–4 indicates the behavior of the four velocities in the neighborhood of an absorption frequency ω_0.

---

* The method is to make a Fourier analysis of the incident pulse, let each component be propagated with the phase velocity $v(\omega)$, and resynthesize the pulse.

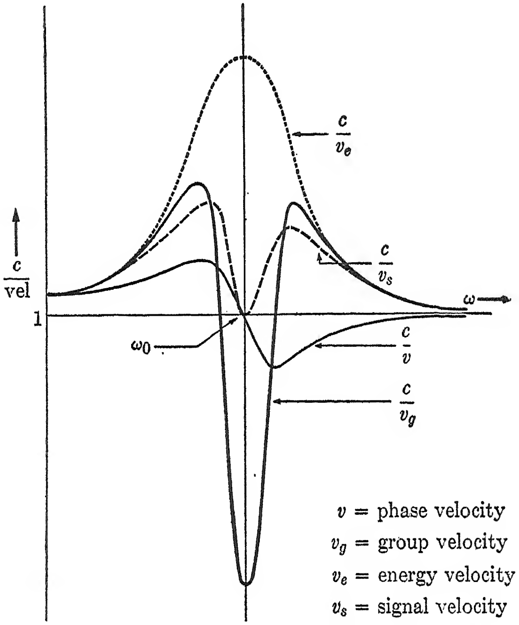


FIG. 21-4 Behavior of various velocities near a resonance frequency (after Brillouin).

21-7 Scattering from a volume distribution. Rayleigh scattering. That part of the scattered radiation coherent with the incoming wave and in the same direction has been successfully treated by means of the total polarization of the assemblage. For the scattering at any other angle it is, in general, necessary to superpose the individual wavelets with the proper phases. If the electrons are distributed completely at random it may be assumed that the phase differences are as often positive as negative and thus average out, so that the total scattered intensity is simply the sum of the individual scattered intensities. If the scatterers are regularly placed, as in a crystal, this is no longer true, and the effect is precisely that of a three-dimensional diffraction grating: in any direction for which successive contributions from neighboring scattering centers are in phase there is a maximum of intensity. For x-rays in ordinary crystals there may be many such directions, giving rise to what is known as Bragg diffraction.* If the wavelength is long in comparison with the distance between neighbors, however, the condition for a maximum cannot be fulfilled except in the forward direction, just as a one-dimensional diffraction grating does not produce a pattern if the grating space is much smaller than λ.

---

* Actually the three-dimensional nature of a crystal "grating" has the consequence that constructive interference is obtained only at a given wavelength *and* a given angle from any set of planes containing regularly spaced scatterers. Thus a single crystal must be rotated during the exposure to monochromatic x-rays if all the regularities are to be evidenced in the diffraction pattern. In a one- or two-dimensional grating there is always an angle at which constructive interference at a given wavelength is obtained if the wavelength is sufficiently short.

In connection with this last point, there arises a question about random scatterers. Consider a cubical volume whose edge is half a wavelength, and which contains many atoms or molecules. Even though the electrons may be distributed randomly within the cube, its radiation will be on the average out of phase with that from a similar neighboring volume in a line of observation at right angles to the beam. The net observed scattering from many such volumes should then be zero from any angle, except, of course, the straightforward direction.* Now these are the conditions on atmospheric scattering of visible light—why, then, is Rayleigh's result in accord with observation?

The answer to this question has been given by Einstein and Smolukowski on the basis of a study of density fluctuations in the atmosphere. The number of scatterers per "cell," or per cube of dimension $\lambda/2$, is constant only on the average, and the fluctuations in this number are responsible for the scattering. If N_j is the number of scattering electrons per cell, and δN_j is the difference of the actual number from the mean, the scattering intensity is proportional to

$$\left|\sum_j e^{i\varphi_j}\,\delta N_j\right|^2 = \sum_j (\delta N_j)^2 + \sum_{j\neq k} e^{i(\varphi_j-\varphi_k)}\,\delta N_j\,\delta N_k, \qquad (21\text{–}48)$$

where φ_j is the phase of the jth cell. To make a comparison with observation, we must take the average of Eq. (21–48). In dilute gases the fluctuations of different cells are independent, and since the average of δN_j is zero, by definition, the second term on the right cancels out. If the fluctuations are small the "dispersion," the square of what is called in probability theory the root-mean-square deviation, is given by

$$\overline{(\delta N_j)^2} = N_j. \qquad (21\text{–}49)$$

Therefore the total scattering is simply proportional to $\Sigma N_j = N$, the total number of scattering centers, in agreement with Rayleigh's assumption. For denser media, the evaluation of the average of Eq. (21–48) is more complicated, and near a chemical phase change these fluctuations give rise to what is called critical opalescence.

21–8 Absorption of radiation by an oscillator.

The consequences of Eq. (21–19) have thus far been discussed principally in terms of the effect of a bound electron on the propagation and scattering of electromagnetic radiation. Another question of fundamental interest concerns the amount of energy abstracted from the wave and converted to energy of the bound electron system.

---

* Incidentally, this argument justifies the procedure used in Section 21–5 for treating dispersive media; if fluctuations are neglected we must consider the effect of the individual oscillations only along the original direction of propagation.

Consider a somewhat more complicated case than a monochromatic plane wave. Let E_ω represent the Fourier component of a general time dependent $E(t)$, i.e., $E = \int E_\omega e^{-i\omega t}\, d\omega$, and let x_ω be the Fourier component of the resultant displacement $x(t)$. Clearly, as with a monochromatic wave,

$$x_\omega = \frac{eE_\omega}{m} \frac{1}{\omega_0^2 - \omega^2 - i\omega\gamma}. \tag{21–50}$$

Let us now consider the rate of work done by the electric field on the electron, and compute the time integral of that work. To make this computation, we must generalize somewhat the theorem of Eq. (13–36). In Section 13–3 we proved that if a real variable, $A(t)$, is represented by a Fourier integral,

$$A(t) = \int_{-\infty}^{\infty} A_\omega e^{-i\omega t}\, d\omega,$$

then

$$\int_{-\infty}^{\infty} A^2(t)\, dt = 4\pi \int_0^{\infty} |A_\omega|^2\, d\omega. \tag{21–51}$$

The proof of this theorem can be readily extended to the product of two real time-dependent variables $A(t)$ and $B(t)$ to give

$$\int_{-\infty}^{\infty} A(t)B(t)\, dt = 2\pi \int_0^{\infty} (A_\omega B_\omega^* + A_\omega^* B_\omega)\, d\omega. \tag{21–52}$$

With this theorem, it follows from Eq. (21–50) that the energy ΔU absorbed by the electron due to the passage of the electric field pulse is

$$\Delta U = \int_{-\infty}^{\infty} eE(t)\dot{x}(t)\, dt = \frac{2\pi e^2}{m} \int_0^{\infty} |E_\omega|^2 \frac{2\omega^2\gamma}{(\omega_0^2 - \omega^2)^2 + \omega^2\gamma^2}\, d\omega. \tag{21–53}$$

This integral can be evaluated under the assumption that the resonance term is sharply peaked relative to the frequency distribution of the external field. In that case,

$$\Delta U = \frac{2\pi e^2}{m} |E_{\omega_0}|^2\, 2\omega_0^2\gamma \int_{-\omega_0}^{\infty} \frac{d\xi}{(2\omega_0\xi)^2 + \omega_0^2\gamma^2}, \tag{21–54}$$

where $\xi = \omega - \omega_0$. We may replace the lower limit by $-\infty$ to obtain the simple result

$$\Delta U = \frac{2\pi^2 e^2}{m} |E_{\omega_0}|^2. \tag{21–55}$$

Note that the result is independent of γ; the energy absorbed from an incident field satisfying the assumption above is thus independent of line width, regardless of the physical cause of such a width.

The result stated in Eq. (21–55) can be brought into a different form. Let S be the energy flow per unit area in the incident beam, which is given by

$$S = \int_{-\infty}^{\infty} \epsilon_0 c E^2 \, dt = 4\pi\epsilon_0 c \int_0^{\infty} |E_\omega|^2 \, d\omega, \tag{21–56}$$

according to Eq. (13–36) or Eq. (21–51). Hence, if we define S_ω by

$$S = \int_0^{\infty} S_\omega \, d\omega, \tag{21–57}$$

Eq. (21–55) becomes simply, in terms of the classical electron radius r_0,

$$\Delta U = 2\pi^2 r_0 c S_{\omega_0}. \tag{21–58}$$

Then, since ΔU represents the energy abstracted from a given energy flow S per unit area, we can write Eq. (21–58) as an integral of the absorption cross section over the frequency:

$$\int_0^{\infty} \sigma(\omega) \, d\omega = 2\pi^2 r_0 c. \tag{21–59}$$

This is the classical "sum rule," giving the total absorption cross section integrated over the frequency. The presence of more than one binding frequency for a single charge will not modify Eq. (21–59).

Some further remarks are appropriate here. Equation (21–59) gives the integrated cross section for "dipole absorption" only, since the "excursion" x of the oscillator is taken to be small compared with the wavelength. Note also that we assumed the displacement $\mathbf{x}$ to be parallel to $\mathbf{E}$. This would be the case if the incident radiation is polarized and directed onto a one-dimensional oscillator so that $\mathbf{E}$ is parallel to $\mathbf{x}$, or if an arbitrary wave group is incident on a three-dimensional oscillator. If a randomly directed and polarized radiation field is incident on a one-dimensional oscillator, then Eqs. (21–55), (21–58), and (21–59) should be divided by 3, because Eq. (21–53) would then contain a factor $\cos^2\theta$, where θ is the angle between $\mathbf{x}$ and $\mathbf{E}$, and the mean value of $\cos^2\theta$ averaged over a sphere is $\frac{1}{3}$.

Equation (21–59) can be generalized to any mechanism of electric dipole absorption. We must remember, however, that if this equation is applied to a system where the charged particles are not light compared with the total mass of the system, e must be replaced by $e_{\text{effective}}$, where

$$\frac{e_{\text{eff}}}{e} = \frac{\text{displacement of charge relative to center of mass}}{\text{absolute displacement of charge}}$$

AGRICULTURAL

and m becomes the "reduced" mass of the absorbing system. If more than one charged particle is present, the sum rule can be generalized to

$$\int_0^\infty \sigma(\omega)\, d\omega = 2\pi^2 c \sum_i \frac{e_{i\,\mathrm{eff}}^2}{4\pi\epsilon_0 m_i c^2}. \tag{21-60}$$

21-9 Equilibrium between an oscillator and a radiation field. In the previous section, we obtained an expression for the energy absorbed by an elastically bound electron in an external energy flux S of radiation whose spectral distribution is $S_\omega\, d\omega$. Let us now consider a linear oscillator in a randomly directed and polarized radiation field whose spectral distribution of energy density U_v is given by

$$U_v = \int_0^\infty U_{v\omega}\, d\omega. \tag{21-61}$$

Let ΔU now be the energy absorbed in a given time Δt. Clearly, $U_{v\omega} = S_\omega/c\, \Delta t$, and in order to make use of Eq. (21-58) we must divide the right side by 3 for reasons discussed above. Hence,

$$\frac{\Delta U}{\Delta t} = \frac{2\pi^2}{3} r_0 c^2 U_{v\omega}. \tag{21-62}$$

From Eq. (19-19) we know that the energy radiated by such an oscillator in a time Δt is given by

$$\frac{\Delta U_{\mathrm{rad}}}{\Delta t} = \frac{2}{3} r_0 \frac{m\ddot{u}^2}{c}. \tag{21-63}$$

Averaged over a cycle of the oscillation, Eq. (21-63) becomes

$$\frac{\overline{\Delta U_{\mathrm{rad}}}}{\Delta t} = \frac{2}{3} \frac{r_0 \omega^2}{c} \overline{U}, \tag{21-64}$$

since $m\overline{\dot{u}^2} = m\overline{\dot{u}_0^2}/2 = \overline{U}$, where $\dot{u}_0$ is the crest value of $\dot{u}$ and where $\overline{U}$ is the energy of the oscillator. Now consider the oscillator to be surrounded by a wall (at the same temperature as the oscillator) which absorbs the emitted radiation and re-emits it randomly. Under these conditions, the oscillator will establish equilibrium with its surroundings, and hence Eqs. (21-64) and (21-62) can be equated. Thus

$$U_{v\omega} = \frac{\omega^2}{\pi^2 c^3} \overline{U}. \tag{21-65}$$

This relation governs the frequency distribution of the energy density of electromagnetic radiation in a cavity with perfectly "black" walls; the distribution is controlled by the statistical behavior of the oscillator at a given temperature. It is this conclusion which leads to the paradoxical Rayleigh-Jeans spectral distribution of blackbody radiation if classical statistical mechanics is applied to the oscillator, i.e., if we put $\bar{U} = kT$. Obviously, the energy density corresponding to a particular frequency cannot increase indefinitely with increasing frequency at a given temperature.

The distribution law, Eq. (21–65), can be arrived at in an entirely different manner. It will be shown in Chapter 24 that the field equations of a bounded radiation field can be transformed in such a way that they are equivalent to the equations for simple harmonic oscillators. In fact, there is a one-to-one correspondence between equivalent oscillators and the normal modes of the field in the enclosure. The coefficient ω^2/π^2c^3 in Eq. (21–65) then corresponds to the number of normal modes per unit circular frequency interval (compare with the answer to problem 8 in the exercises at the end of Chapter 12).

SUGGESTED REFERENCES

R. Becker, *Theorie der Elektrizität,* Band II. Sections B, on the elastically bound electron, and C III, on dispersion and absorption, are highly recommended.

L. Landau and E. Lifshitz, *The Classical Theory of Fields.* An excellent account of radiation and scattering will be found in Chapter 9.

H. A. Lorentz, *Problems of Modern Physics.* Many of the subjects covered in this series of lectures given in 1922 are still of great interest, but the work is cited here because it includes the application of Rayleigh's formula to the atmosphere.

L. Rosenfeld, *Theory of Electrons.* Probably the best account from a modern point of view of the classical theory of dispersion, including critical opalescence.

J. A. Stratton, *Electromagnetic Theory.* Includes a treatment of ionosphere transmission.

EXERCISES

1. Show that the scattering cross section of a free electron for an elliptically polarized wave (electric field $\mathbf{E} = \mathbf{A}\cos\omega t + \mathbf{B}\sin\omega t$, where $\mathbf{A}$ and $\mathbf{B}$ are mutually perpendicular vector amplitudes) is given by

$$\frac{d\sigma}{d\Omega} = r_0^2\left[\frac{(\mathbf{A}\times\mathbf{n})^2 + (\mathbf{B}\times\mathbf{n})^2}{A^2 + B^2}\right],$$

where $\mathbf{n}$ is a unit vector in the direction of the radiation.

2. What is the minimum frequency that can be propagated in free space without attenuation, assuming one electron per cubic centimeter? Find the group velocity for higher frequencies.

3. Show that in the presence of a magnetic field $\mathbf{H}$ in the direction of wave propagation an electron gas exhibits two velocities with indexes of refraction given by

$$n_\pm^2 = 1 - \frac{Ne^2/m\epsilon_0\omega^2}{1 \pm \dfrac{e\mu_0 H}{m\omega}} = 1 - \frac{4r_0N/k^2}{1 \pm \omega_L/\omega},$$

where ω_L is the Larmor frequency.

4. Make a sufficiently accurate plot of the wave and group velocities and the absorption coefficient near a resonance frequency to indicate the essential features, showing that the maximum and minimum of the index of refraction occur at frequencies where the absorption coefficient has half its maximum value.

5. Compute the rate of energy transport, i.e., the time average of the Poynting vector over that of the energy density, at $\omega = \omega_0$.

6. How much polarization is to be expected in sky light scattered at angle θ? Assume unpolarized light from the sun and take account only of single scattering.

7. In Eq. (21–43) take $a(k) = e^{-(k-k_1)^2b^2/2}$, with k_1 and b constants. Find the form of the group in space, and trace it in time as it progresses into a dielectric. Find a general expression for its width in space, i.e., the distance from the position of maximum amplitude to the place where the amplitude is $1/e$ of its maximum value. Assume that the frequency of the group is not too near a resonance frequency of the dielectric.

8. Prove Eq. (21–52).

9. Consider a nucleus of N neutrons and Z protons, and let $A = N + Z$. Show that under the assumption of pure electric dipole absorption, the "integrated cross section" is given by

$$\int \sigma \, dE = 2\pi^2 \frac{NZ}{A} \alpha \left(\frac{\hbar}{Mc}\right)^2 Mc^2,$$

where α is the fine-structure constant, $\simeq 1/137$, and M is the proton mass, approximately equal to that of the neutron. Evaluate this expression numerically.

CHAPTER 22

COVARIANT FORMULATION OF THE FIELD EQUATIONS IN MATERIAL MEDIA AND THE CONSERVATION LAWS OF ELECTRODYNAMICS

From the electron theory viewpoint the ordinary Maxwell equations are the result of averaging electrical quantities over regions large enough to permit macroscopic observation. We saw in Chapter 17 that vacuum electrodynamics with charges and currents may be described in a simple covariant manner. It should then be possible to write the equations of electrodynamics in material media in covariant form.

22–1 Covariant description of sources. The new element in macroscopic electrodynamics of material media is that the four-vector current j^i may have four nonvanishing components even in a frame in which the medium is at rest. In such a frame,

$$(j^i) = \left(\frac{\mathbf{j}^0}{c}, \rho^0\right), \tag{22–1}$$

where $\mathbf{j}^0$ is the current density in the proper frame. Let us now consider the form of the components of j^i in a nonproper frame. To correspond to the vacuum case, j^i must retain the components

$$(j^i) = \left(\frac{\mathbf{j}}{c}, \rho\right) \tag{22–2}$$

in any frame, and hence, in general,

$$j_x = \gamma(j_x^0 - \rho^0 v), \tag{22–3}$$

$$\rho = \gamma\left(\rho^0 - \frac{j_x^0 v}{c^2}\right), \tag{22–4}$$

where v is the velocity (taken along x) of the arbitrary frame measured in the proper frame.

Equation (22–3) is physically clear: it contains the convective current due to transport of charge and the contraction factor which assures the invariance of the charge. Equation (22–4) is less obvious physically: it says that a substance which carries a current but is electrically neutral ($\rho^0 = 0$) in a proper frame does not necessarily remain neutral when observed from another inertial frame. This effect can be understood in terms

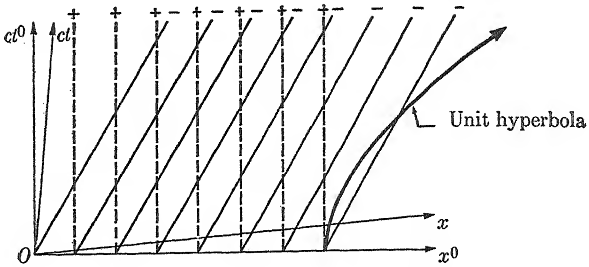


FIG. 22–1 Minkowski diagram showing the possibility of unequal charge densities in two Lorentz frames.

of the kinematics of the moving charges. Let us consider equal densities of positive and negative charges in the proper frame Σ^0. For simplicity, let the positive charges be at rest (although this is not necessary to the argument) and the negative charges be in motion. The world lines of the plus charges (shown as dashes) and of the minus charges (shown solid) as observed in Σ^0 and in an arbitrary frame Σ are indicated in the Minkowski diagram, Fig. 22–1. The charge density in Σ^0 is measured by counting + and − charges simultaneously in Σ^0, i.e., on the x^0-axis; the charge density in Σ is measured by counting + and − charges along the x-axis, i.e., averaging them simultaneously in Σ. It is seen that the density of − charges along the x-axis is decidedly decreased relative to that along the x^0-axis, while the density of + charges has changed very little. (Note, however, that density of charge is measured by counting charges per unit length defined by the intercept of the unit hyperbola with the respective axes, as shown in the figure.) Therefore a net positive charge is found in Σ, corresponding to the negative current in the neutral proper frame, in agreement with Eq. (22–4).

One consequence of this effect, resulting directly from the difference in simultaneity between Σ^0 and Σ, is the fact that a neutral stationary current loop in Σ^0 acquires an electric moment when observed in Σ. Consider the rectangular current loop of Fig. 22–2, carrying a current J. From

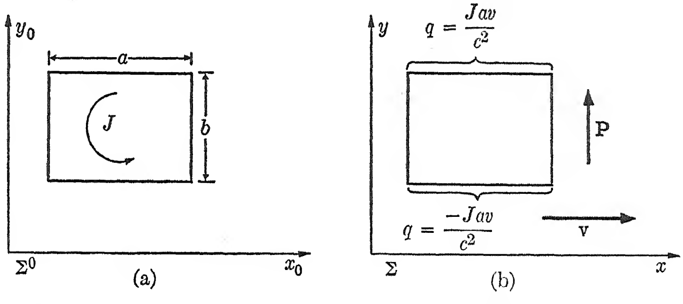


FIG. 22–2 Electric moment acquired by a current loop when viewed from a moving frame.

the point of view of Σ the legs parallel to x^0 will carry charges of $\pm Jav/c^2$ respectively, if we ignore the terms quadratic in v. Thus the system has an electric moment

$$P = \frac{vabJ}{c^2} = -\frac{|\mathbf{v} \times \mathbf{m}|}{c^2}, \tag{22-5}$$

where $\mathbf{m} = J\mathbf{S}$ is the magnetic moment of the loop of vector area $\mathbf{S}$, $|\mathbf{S}| = ab$.

22-2 The field equations in a material medium. If we divide the charge-current four-vector into "true" components j^i and magnetization-polarization components j^i_M, then Maxwell's source equations, Eq. (17-34), become

$$\frac{\partial F^{il}}{\partial x^i} = \frac{j^l + j^l_M}{\epsilon_0}, \tag{22-6}$$

while Eq. (17-36) remains unchanged. As before, the F^{ij} are given by the matrix of Eq. (17-32) in terms of $\mathbf{B}$ and $\mathbf{E}$. It is desirable, as in the three-dimensional formulation, to write Eq. (22-6) in terms of the true charges and currents as external sources only, and to incorporate the "induced" charges and currents into the fields. This can be done by introducing the moment tensor M^{ij} by the equation

$$j^l_M = \frac{\partial M^{il}}{\partial x^i}. \tag{22-7}$$

In order to correspond to the inaccessible sources of Chapter 7, the components of j^i_M must be

$$j^l_M = \left(\frac{\mathbf{j}_M + \mathbf{j}_P}{c}, \rho_P\right) = \left(\frac{\nabla \times \mathbf{M}}{c} + \frac{1}{c}\frac{\partial \mathbf{P}}{\partial t}, -\nabla \cdot \mathbf{P}\right). \tag{22-8}$$

The components of M^{ij} are then given by

$$(M^{ij}) = \begin{array}{c} \\ i \\ \downarrow \\ \\ \end{array} \begin{array}{c} j \rightarrow \\ \left(\begin{array}{cccc} 0 & -M_z/c & M_y/c & -P_x \\ M_z/c & 0 & -M_x/c & -P_y \\ -M_y/c & M_x/c & 0 & -P_z \\ P_x & P_y & P_z & 0 \end{array}\right) \end{array}. \tag{22-9}$$

Corresponding to the three-dimensional relations

$$\mathbf{H}/c\epsilon_0 = c\mathbf{B} - \mathbf{M}/c\epsilon_0, \tag{22-10}$$

$$\mathbf{D}/\epsilon_0 = \mathbf{E} + \mathbf{P}/\epsilon_0, \tag{22-11}$$

we may introduce a new field H^{ij} defined by

$$H^{ij}/\epsilon_0 = F^{ij} - M^{ij}/\epsilon_0. \tag{22–12}$$

The source equations, Eq. (22–6), then become simply

$$\frac{\partial H^{ij}}{\partial x^i} = j^j, \tag{22–13}$$

where

$$(H^{ij}) = \begin{array}{cc} & j \rightarrow \\ \begin{array}{c} i \\ \downarrow \end{array} & \begin{pmatrix} 0 & -H_z/c & H_y/c & D_x \\ H_z/c & 0 & -H_x/c & D_y \\ -H_y/c & H_x/c & 0 & D_z \\ -D_x & -D_y & -D_z & 0 \end{pmatrix} \end{array} . \tag{22–14}$$

The signs in the defining equations of the auxiliary fields correspond to the way in which the equivalent currents and charges are derived from the moments.

22–3 Transformation properties of the partial fields. The transformation properties of the moments follow directly from the covariant formulation, as in Chapter 17 for the case of the total fields. We obtain for the transformation from the plain to the primed system moving with relative velocity v,

$$P'_{\parallel} = P_{\parallel}, \tag{22–15}$$

$$M'_{\parallel} = M_{\parallel}, \tag{22–16}$$

$$\mathbf{P}'_{\perp} = \gamma\left(\mathbf{P}_{\perp} - \frac{\mathbf{v} \times \mathbf{M}_{\perp}}{c^2}\right), \tag{22–17}$$

$$\mathbf{M}'_{\perp} = \gamma(\mathbf{M}_{\perp} + \mathbf{v} \times \mathbf{P}_{\perp}). \tag{22–18}$$

Equation (22–15) is to be expected, since $P_{\parallel}$ is the product of an (invariant) charge and a distance divided by a volume, both contracted in the same ratio. A similar argument applies to Eq. (22–16). The term $\mathbf{v} \times \mathbf{P}$ in Eq. (22–18) is, apart from the factor γ, nonrelativistic, arising from the fact that convection of a polarized medium corresponds to a net circulation of charge. We have met this effect before (Chapter 9) in the discussion of Maxwell's equations in moving media from a nonrelativistic point of view. Consider an infinite polarized slab, shown in Fig. 22–3. From Σ, a moving frame, this slab possesses opposing surface currents corresponding to a uniformly magnetized medium.

Equation (22–17) has no nonrelativistic counterpart. It represents, as already shown, the effect of the net charges when a current in a neutral

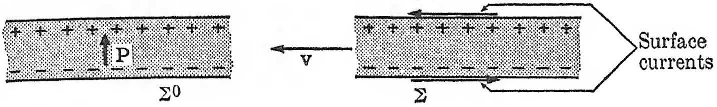


FIG. 22-3 A polarized slab in motion corresponds to a magnetized medium.

stationary conductor is viewed from a nonproper frame. The extra electric moment is that predicted by Eq. (22-5), and arises from the relativistic definition of simultaneity. This equivalent electric moment resolves the apparent paradox of the unipolar induction generator. We concluded in Section 9-5 that a current would flow when a conductor, in contact with take-off brushes, moves transverse to a magnetic field. No difficulties arise so long as the source of **B** is external; the effect should, however, persist if **B** is due to a permanent magnetization of the bar itself. Since the permanent magnetization is to be describable in terms of equivalent Amperian currents alone, the question arose as to how such currents could produce an electrostatic effect when viewed from a moving frame. Now we see that this description leads to no difficulties, since we can interpret the electrostatic field as due to the equivalent moment $\mathbf{v} \times \mathbf{M}/c^2$. Inasmuch as this equivalent moment is a consequence of the relativistic definition of simultaneity, unipolar induction is fundamentally a relativistic effect.

Practical devices for producing unipolar induction ordinarily involve rotation, and consequently some remarks should be added here concerning rotary motion or accelerated motion in general. A full treatment of accelerated frames is beyond the scope of this book, but some consideration is necessary to clarify the applicability of statements made here and in Section 9-5 to actual experiments.

Contrary to what is true of frames differing in motion only by uniform relative velocity (Lorentz frames), experiments *can* be performed that distinguish rotating (accelerated) frames. Here we have been dealing with the special (inertial) frames in which the laws of special relativity are valid. Which frame is an inertial frame is presumably defined by gravitational forces, and therefore "a preferred frame as to rotation" is defined by the location of the bulk of the masses in the universe. Strictly speaking, there is no frame in which special relativity is exactly valid, since no frame can be found in which gravitational accelerations or "equivalent" (i.e., indistinguishable) inertial accelerations vanish. For all electrical phenomena, however, the surface of the earth is a satisfactory Lorentz frame to a very good approximation; note that the tests of the special theory are concerned with the effects of the linear, not the curvilinear, motion of the earth's surface.

But there are large scale electromagnetic effects which distinguish noninertial frames produced by rotation of matter in the earth's frame. Many paradoxes result if one assumes that such phenomena should be reciprocal in the rotating frame and that of the earth. Let us consider the device analogous in rotation to the sliding-bar experiment of Section 9–5, the "Faraday disk" illustrated in Fig. 22–4. Its behavior conforms in some details to the conclusions drawn in Section 9–5. All the results set forth in Table 9–1, which gives the electromotive force corresponding to various possibilities of motion of the disk, the source of magnetic field, and the observing apparatus, are preserved. In particular, the important conclusion is retained that motion (rotation in this case) of the source of magnetic field does not affect any physical process, so long as such motion does not produce a time-varying field. Again the electromotive force can be calculated directly from Faraday's law, Eq. (9–2), if the change in flux is computed through an orbit which moves with the actual carrier of the current; here this path is split between the two frames. Also, consideration of the "effective force on an electron" gives the right answer. Note, however, that the effective electric field $\mathbf{E}_r$ in the rotating frame is given by $\mathbf{E}_r = \omega B r$ for low velocities; this expression has a nonvanishing divergence, and thus a volume charge is developed. This is at variance with our transformation laws for linear motion, and is an indication that the "absolute" rotational motion of the disk (i.e., the motion relative to an inertial frame) can in principle be determined.

Another experiment which more obviously distinguishes absolute rotational frames is that shown in Fig. 22–5. A negative point charge is surrounded by a positive spherical surface distribution of equal total charge. If this system is not rotating, $\mathbf{E} = \mathbf{B} = 0$ outside the sphere, in accord with Gauss's theorem. But if the charge is rotating the positive charges constitute ring currents, and thus the entire assembly develops a magnetic moment. Hence $\mathbf{B}$ is different from zero, both inside and outside the sphere. Again our transformation laws fail to apply to nonrectilinear

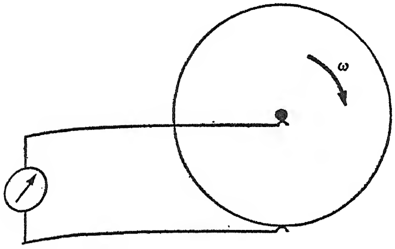


FIG. 22–4 Faraday disk illustrating unipolar induction in rotation.

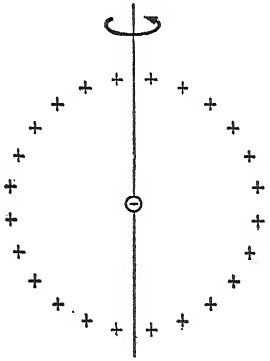


FIG. 22–5 Rotating charge distribution of zero net charge.

motion. It is not possible to discuss the problem consistently in both the stationary and the rotating frames without recourse to general relativity.*

It should be emphasized, however, that within the framework of special relativity our earlier considerations do permit us to describe all fields in all Lorentz frames, whether the sources are accelerated or not.

22–4 Covariant formulation of the conservation laws of vacuum electrodynamics. The conservation laws of Chapter 10 may be written in three-dimensional tensor language as

$$\frac{\partial N_\alpha}{\partial x_\alpha} + \rho E_\alpha u_\alpha + \epsilon_0 \frac{\partial}{\partial t}\left[\frac{(E^2 + c^2B^2)}{2}\right] = 0, \tag{22–19}$$

$$\frac{\partial T^\circ_{\alpha\beta}}{\partial x_\beta} - \epsilon_0 \frac{\partial}{\partial t}(\mathbf{E} \times c\mathbf{B})_\alpha = \rho[E_\alpha + (\mathbf{u} \times \mathbf{B})_\alpha]. \tag{22–20}$$

Here $\mathbf{N} = \mathbf{E} \times \mathbf{H}$ is the Poynting vector and $T^M_{\alpha\beta} = E_\alpha D_\beta + H_\alpha B_\beta - \delta_{\alpha\beta}(E_\gamma D_\gamma + H_\gamma B_\gamma)/2$ is the Maxwell stress tensor. Equation (22–19) balances energy of radiation loss with the rate of mechanical and thermal work and the time rate of change of the electromagnetic field energy. Equation (22–20) balances the volume force with the rate of increase of the mechanical and electromagnetic momentum. These two equations can be combined into a single relation by introducing the energy-momentum tensor of the electromagnetic field defined formally by the symmetric matrix

$$T^{ij} = \begin{pmatrix} -T^M_{11} & -T^M_{12} & -T^M_{13} & cG_x \\ -T^M_{12} & -T^M_{22} & -T^M_{23} & cG_y \\ -T^M_{13} & -T^M_{23} & -T^M_{33} & cG_z \\ cG_x & cG_y & cG_z & W \end{pmatrix}, \tag{22–21}$$

where $\mathbf{G} = \mathbf{N}/c^2$ is the momentum density † of the field and $W = (\epsilon_0 E^2 + \mu_0 H^2)/2$ is the energy density. If T^{ij} is a tensor with the components as given, the conservation laws, Eqs. (22–19) and (22–20), are equivalent to the simple covariant relation

$$\frac{\partial T^{ij}}{\partial x^j} = -f^j, \tag{22–22}$$

where

$$(f^i) = \rho(\mathbf{E} + \mathbf{u} \times \mathbf{B}, \mathbf{E} \cdot \mathbf{u}/c) \tag{22–23}$$

is the four-vector representing the Lorentz force per unit volume and the rate of work per unit volume of the electromagnetic field on the charges and currents.

---

* See L. I. Schiff, *Proc. Nat. Acad. Sci.* **25,** 391 (1939).

† The relation $\mathbf{G} = \mathbf{N}/c^2$ can be taken as a consequence of the symmetry of the tensor $T^{ij} = T^{ji}$.

To show that T^{ij} is actually a tensor, we note that it can be generated from F^{ij} of Eq. (17–32) by the tensor operation

$$T^{ij} = \epsilon_0(F^{ik}F^j_k - \tfrac{1}{4}\,\delta^{ij}F^{kl}F_{kl}), \tag{22–24}$$

where $F^j_k = g_{kn}F^{jn}$. Therefore T^{ij} is a tensor. The algebraic correctness of Eq. (22–22) can then be easily checked. It is seen that the second term on the right side of Eq. (22–24) is simply the invariant "trace" of F^{kl}, namely,

$$F^{kl}F_{kl} = 2(c^2B^2 - E^2). \tag{22–25}$$

The invariance of $c^2B^2 - E^2$ again shows that the ratio between the electric and magnetic fields in a plane electromagnetic wave in space is a constant.

22–5 Some consequences of covariant formulation of the conservation laws. The conservation laws, like the field equations, are in agreement with relativity theory without modification. Moreover, a Lorentz transformation of Eq. (22–21) again shows the equivalence of energy transport and flow of momentum. A number of other interesting and important conclusions can be drawn from the covariant conservation laws and the form of the energy momentum tensor.

1. If "state of equilibrium" is to be an invariant property of a system, we must conclude not only that electromagnetic forces, energy, momenta, etc., must be describable by a tensor relation of the form $f^j = -\partial T^{ij}/\partial x^i$, but also that the totality of such mechanical quantities must obey an equation of this form. Hence, for example, Eq. (22–22) will be a valid equation, in a medium under elastic stress, with W as the mass density (including elastic energy density), G the mechanical momentum density, and $\overset{\circ}{T}_{\alpha\beta}$ the elastic stresses. This formulation yields the transformation equations for all quantities entering into the mechanics of continua. Note particularly that the mass density of a continuous medium cannot be treated as a scalar, or even as a component of a four-vector like the electrical charge density, but is the (4,4) component of the mechanical energy momentum tensor. The quantity $m = \int W\,dv$ has the same transformation character as a point mass. These facts are consistent with the existence of a fundamental unit of charge and the apparent lack of existence of a fundamental unit of mass. The tensor component character of mass density is of importance in the formulation of the gravitational action of matter in the general theory of relativity.

2. Let us consider a volume v containing *totally* a quantity of free electromagnetic radiation but *no* charges or currents. The energy tensor of Eq. (22–21) then obeys the conservation law

$$\frac{\partial T^{ij}}{\partial x^i} = 0, \tag{16–24}$$

and hence, according to the theorem of Eq. (16–25),

$$G^i = \left(\int c\mathbf{G}\, dv, \int W\, dv\right) = -\int T^{4i}\, dv \tag{22–26}$$

is a contravariant four-vector. Thus the momentum and energy of a radiation pulse totally contained within a finite volume has the same transformation properties as a material point particle. For the *total* field of a charge this is not true, since Eq. (16–24) is not satisfied; we shall return to this point shortly. For a plane electromagnetic wave the invariant

$$G^iG_i = W^2 - c^2G^2 \tag{22–27}$$

is zero, and hence the equivalent particle properties of such a wave correspond to zero rest mass. This is in agreement with the fact that radiation, propagated with velocity c, can obey the particle transformation laws and still yield finite momenta and energy, which would be impossible if the rest mass were other than zero. All these facts are consistent with the "light quantum" concept.

3. The "phase" of an electromagnetic wave is defined by the relation

$$E = E_0e^{i\varphi} = E_0e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \tag{22–28}$$

where $\mathbf{k}$ is the wave propagation vector and ω is the angular frequency. The zero point of a field must be an invariant physical fact, and hence one would expect φ to be an invariant. Therefore we can write

$$\varphi = k_ix^i \tag{22–29}$$

with

$$k_i = (\mathbf{k}, -\omega/c), \quad k^i = -(\mathbf{k}, \omega/c). \tag{22–30}$$

Since $k_ik^i = 0$, k^i transforms exactly like an energy momentum vector of a particle of zero rest mass. This suggests making the momentum and energy of a light quantum proportional to $\mathbf{k}$ and ω respectively, as has been done in quantum theory.

Equation (22–30), in defining the transformation character of $\mathbf{k}$ and ω, provides a simple method of obtaining the relativistically correct expressions for the Doppler shift and for the aberration of starlight. Consider a source at rest in a frame Σ^0 radiating in the x-direction such that $k_x^0 = \omega^0/c$. In a Σ frame we obtain from the Lorentz transformation applied to k^4:

$$\omega = c\gamma\left(\frac{\omega^0}{c} - \beta k_x^0\right) = \frac{\omega^0(1-\beta)}{\sqrt{1-\beta^2}} = \omega^0\sqrt{\frac{1-\beta}{1+\beta}}. \tag{22–31}$$

This is the relativistic formula for the Doppler shift. The expression for aberration can be obtained similarly from the spatial components of k^i.

The relation

$$\sin\theta = \beta \tag{22-32}$$

is obtained for the aberration angle θ, in agreement with Eq. (15–48).

4. It is to be noted that only the electromagnetic momentum and energy of a *free* wave have the correct transformation character for a particle of zero rest mass. Let us consider the electromagnetic field of a point charge moving uniformly with velocity $\mathbf{u}$ in the x-direction. In the Lorentz frame in which the charge is at rest, the energy momentum tensor would have the components

$$T^{ij} = \begin{pmatrix} -T^{M_0}_{\alpha\beta} & 0 \\ 0 & W_0 \end{pmatrix}, \tag{22-33}$$

to be integrated over a three-dimensional volume; W_0 is the electrostatic field energy density of the charge. In the general frame, it follows from the Lorentz transformation that

$$T^{14} = \frac{\beta(W_0 - T^{M_0}_{11})}{1-\beta^2}, \tag{22-34}$$

$$T^{44} = \frac{W_0 - \beta^2 T^{M_0}_{11}}{1-\beta^2}. \tag{22-35}$$

Note that we chose the sense of the Lorentz transformation such that frame Σ^0 has a positive velocity in Σ. Hence the momentum and energy of the field become

$$\begin{aligned} cG_x &= \int T^{14}\,dv = \int \sqrt{1-\beta^2}\,T^{14}\,dv_0 \\ &= \frac{\beta}{\sqrt{1-\beta^2}} \int (W_0 - T^{M_0}_{11})\,dv_0, \end{aligned} \tag{22-36}$$

$$\begin{aligned} U &= \int T^{44}\,dv = \int \sqrt{1-\beta^2}\,T^{44}\,dv_0 \\ &= \frac{1}{\sqrt{1-\beta^2}} \int (W_0 - \beta^2 T^{M_0}_{11})\,dv_0, \end{aligned} \tag{22-37}$$

where dv_0 is the volume element in the rest frame. But if the charge is spherically symmetrical in the rest frame,

$$\int W_0\,dv_0 = U_0 \tag{22-38}$$

and

$$\int T^{M_0}_{11}\,dv_0 = \epsilon_0 \int \left(E_{x0}^2 - \frac{1}{2} E_0^2\right) dv_0 = -\frac{\epsilon_0}{6} \int E_0^2\,dv_0 = -\frac{U_0}{3}, \tag{22-39}$$

where U_0 is the electrostatic energy of the charge, and where we have made use of the relation, satisfied by a spherically symmetrical field, that $E_x^2 = \frac{1}{3}E^2$. Hence,

$$\mathbf{G} = \frac{4}{3}\frac{\mathbf{u}}{c^2}\frac{U_0}{\sqrt{1-\beta^2}}, \tag{22–40}$$

$$U = \frac{U_0}{\sqrt{1-\beta^2}}\left(1+\frac{\beta^2}{3}\right). \tag{22–41}$$

These expressions obviously do not transform as components of a four-vector, and therefore we obtain an additional argument for the fact, discussed in Chapter 20, that the electromagnetic mass of a particle does not constitute its total mass.

22–6 The electromagnetic energy momentum tensor in material media. We have seen that the covariant formulation of the conservation laws for vacuum electrodynamics is both straightforward and fruitful. The separate formulation for the macroscopic fields in material media, however, involves some apparent ambiguities. We shall consider this problem only briefly.

The conservation laws of vacuum electrodynamics, Eq. (22–22), may be written

$$-f^j = \frac{\partial T^{ij}}{\partial x^i} = j_l F^{jl}, \tag{22–42}$$

where j^l is the four-vector current density. If we assume that the volume force f^j in material media is similarly given by $j_l F^{jl}$, where j^l is interpreted as the true (macroscopic) current and charge density satisfying Eq. (22–13), we must introduce a new energy momentum tensor. The equation corresponding to Eq. (22–42) becomes

$$-f^j = \frac{\partial S^{ij}}{\partial x^i}, \tag{22–43}$$

where

$$S^{ij} = g^{jk}(H^{il}F_{kl} - \tfrac{1}{4}\delta^i_k H^{ml}F_{ml}). \tag{22–44}$$

Equation (22–43) is covariant, since S^{ij} as given by Eq. (22–44) is a tensor, and algebraically the conservation laws are correctly given. The only argument against the correctness of this energy momentum tensor is based on the fact that it is not symmetric: $S^{4i} \neq S^{i4}$ for dielectrics and permeable materials. It would seem that if Maxwell's macroscopic equations are indeed just the result of averaging the microscopic vacuum quantities no difference in symmetry properties could arise, and for many years a "corrected" (i.e., symmetrized) energy momentum tensor was preferred by most physicists. More recently, however, it has been pointed out that there is no reason for demanding symmetry in the case of a partial field which also leaves out of consideration all mechanical properties, and that the only valid demand on S^{ij} is that it correspond to the correct Lorentz volume force in all instances. The situation is fully discussed by Møller in the reference cited below, who shows that the evidence tends to favor the formulation of Eq. (22–44), due originally to Minkowski. The total energy momentum tensor, including mechanical as well as electrical quantities, is of course symmetrical, corresponding to a closed system. This follows, since the relation $\mathbf{G} = \mathbf{N}/c^2$ should be obeyed by the total mechanical properties of the system.

SUGGESTED REFERENCES

C. Møller, *The Theory of Relativity.* This work contains an excellent detailed discussion of the conservation laws in general closed systems, and the problems involved in their extension to the partial fields of macroscopic electrodynamics.

R. Becker, *Theorie der Elektrizität.* Abraham's symmetrical energy-momentum tensor, designed to replace the unsymmetrical form derived by Minkowski for electromagnetic media, is presented in this work, in addition to the standard formulation of the macroscopic field tensor.

R. C. Tolman, *Relativity, Thermodynamics, and Cosmology.* Chapter IV is devoted to an exposition of the macroscopic theory of relativistic electrodynamics.

W. Pauli, *Relativitäts Theorie.* This is a classic presentation of the subject, including a thorough discussion of the problems as they appeared in 1921.

EXERCISES

1. Show that the constitutive equations can be written relativistically as

$$H_{ij}u^i = \epsilon F_{ij}u^i,$$

and

$$u_jF_{kl} + u_kF_{lj} + u_lF_{jk} = \mu c^2(u_jH_{kl} + u_kH_{lj} + u_lH_{jk}),$$

where $u_i = dx_i/ds$. Write this in three-dimensional language.

2. Check the algebraic correctness of the conservation laws as derived from the tensor of Eq. (22–24).

3. Find the components of S^{ij} as defined by Eq. (22–24) and show that it implies an electromagnetic momentum density proportional to $\mathbf{D} \times \mathbf{B}$ in material media.

4. Prove that the quantity $F^{ik}F^j_k$ in Eq. (22–24) is symmetrical in i and j.

5. Show that $T^i_i = 0$.

CHAPTER 23

THE MOTION OF CHARGED PARTICLES IN ELECTROMAGNETIC FIELDS

Material media actually consist of charged particles and, as Einstein has pointed out, the energy tensor represents matter only provisionally. The fundamental problem of electromagnetic theory deals with the motion of charged particles interacting via electromagnetic forces. The field concept is not necessary in formulating equations for such motion: formulation in terms of action at a distance is possible in a relativistically covariant way.* Separation of the problem into "radiation" (production of a field by a charge system) and "motion of charged particles in a field" is in a certain sense artificial and introduces some difficulties. In particular, this separation implies that the radiation can be computed from the motion of the system as driven by external forces. In reality, there is also the "radiation reaction" which must be taken into account in order that the conservation laws be satisfied. The reactions can be computed from the radiation and applied as corrections to the external forces for a more accurate description, and this process of successive approximations converges unless the variability of the motion is too high. In case the reactions are negligible, however, the separation is possible and generally advantageous, since the sources of the field need not be known in detail for a description of the motion of a charge. We shall deal here with the motion of charged particles only under the conditions that reactions may be neglected.

23–1 World-line description. In the covariant language of Chapters 16 and 17, the equation of motion for a particle of charge e, rest mass m_0, and four-velocity u^i, moving in an external field F^{ij}, is given by

$$m_0c^2 \frac{du^i}{ds} = eu_jF^{ij}, \tag{23–1}$$

corresponding to the three-dimensional form

$$\frac{d\mathbf{p}}{dt} = e(\mathbf{E} + \mathbf{u} \times \mathbf{B}). \tag{23–2}$$

Here, as before,

$$\mathbf{p} = m_0\mathbf{u}/\sqrt{1 - u^2/c^2}. \tag{23–3}$$

---

* See, e.g., J. A. Wheeler and R. P. Feynman, *Revs. Modern Phys.*, **21**, 425 (1949).

Since Eq. (23–2) is equivalent to the components of Eq. (23–1), it is the relativistic extension of the low velocity equation of motion which follows from the definition of the field.

To find the equations of motion in canonical form, it is customary to proceed by choosing a Lagrangian which yields the correct equations. Here we shall do this in such a way as to emphasize the required invariance properties.

Let the world line of the particle be given by $x^i(s)$, where s is the "proper time" for the particle. Let us look for a Lagrangian L_4 in four-space such that the motion is given by a variational principle,

$$\delta \int_{x_0^i}^{x_1^i} L_4(x^i,u^i)\, ds = 0. \tag{23–4}$$

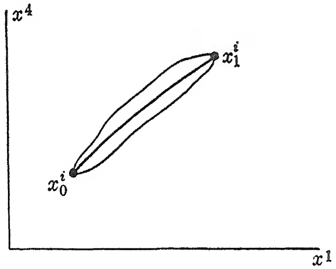


FIG. 23–1 World-line variation of the path of a particle.

The variation is taken between fixed world points x_0^i and x_1^i, as indicated in Fig. 23–1. Invariance requires that L_4 be a scalar, and in order to yield a linear equation of motion, the invariants making up L_4 must be not higher than second order in u^i. For a free particle, we are led to choose

$$L_4 = \frac{m_0c^2}{2}\, u^iu_i, \tag{23–5}$$

explicit dependence on x^i being excluded, since all world points are equivalent for a free particle. (The coefficient of u^iu_i is arbitrary in this case.) We may introduce a scalar interaction with the field described by a four-potential $\phi^i(x^i)$ by putting for the total Lagrangian

$$L_4 = \frac{m_0c^2}{2}\, u^iu_i + eu_i\phi^i. \tag{23–6}$$

[The constants in Eq. (23–6) have been adjusted so that the equations of motion will agree with the usual convention for rest mass and charge.] Let us now apply the variational principle, Eq. (23–4), to obtain the Eulerian equations:

$$\delta \int_{x_0^i}^{x_1^i} L_4(x^i,u^i)\, ds = \int_{x_0^i}^{x_1^i} \left[m_0c^2u_i\, \delta u^i + e\left(\phi_i\, \delta u^i + \frac{\partial \phi_i}{\partial x^j}\, \delta x^j u^i\right)\right] ds. \tag{23–7}$$

Since $\delta u^i = \dfrac{d(\delta x^i)}{ds}$ and $\dfrac{d\phi_i}{ds} = \dfrac{\partial \phi_i}{\partial x^j}\, u^j$, we can integrate by parts to obtain

$$\delta\int_{x_0^i}^{x_1^i} L_4(x^i,u^i)\,ds = \int_{x_0^i}^{x_1^i}\left[-\frac{d}{ds}(m_0c^2u_j) + e\left(\frac{\partial\phi_i}{\partial x^j} - \frac{\partial\phi_j}{\partial x^i}\right)u^i\right]\delta x^j\,ds. \tag{23-8}$$

The integrated part does not appear, since the variation of the function vanishes at x_0^i and x_1^i. Since $\delta x^j(s)$ is arbitrary along the path, the bracketed expression must be equal to zero, and hence

$$m_0c^2\frac{du_j}{ds} = eF_{ji}u^i, \tag{23-9}$$

which is the covariant equivalent of the contravariant Eq. (23-1).

23-2 Hamiltonian formulation and the transition to three-dimensional formalism. A Hamiltonian in the ordinary sense cannot be readily introduced in covariant formulation, since in the ordinary Hamiltonian equations,

$$\frac{\partial H}{\partial p_\alpha} = \frac{dq_\alpha}{dt}, \quad \frac{\partial H}{\partial q_\alpha} = -\frac{dp_\alpha}{dt}, \tag{23-10}$$

the time enters asymmetrically. On the other hand, if we introduce a four-Hamiltonian, H_4, which is to obey the relations

$$\frac{\partial H_4}{\partial p^i} = \frac{dq_i}{ds}, \quad \frac{\partial H_4}{\partial q^i} = -\frac{dp_i}{ds}, \tag{23-11}$$

then H_4 (which is different from H) will be a scalar. To establish a connection between Eq. (23-11) and Eq. (23-10), we must recognize that (23-11) represents eight equations, while (23-10) represents six. We shall show that if we find a scalar $H_4(p^i,q^i)$ such that (23-11) yields the correct equation of the world point, then the magnitude of this scalar is a universal constant for a given particle. If we then solve this equation for $p^4(x_\alpha,p_\alpha,x^4)$, we can show that p^4 obeys Eqs. (23-10) for H. Since p^4 is thus generated by a covariant process, the resultant equations, although not "manifestly" covariant, are nevertheless relativistically correct. We can, of course, add the final test that the correct relativistic equations of motion shall result.

The four-Hamiltonian obeying Eq. (23-11) is defined by

$$H_4 = p^iu_i - L_4, \quad p^i = \frac{\partial L_4}{\partial u_i}, \tag{23-12}$$

or, from Eq. (23-6),

$$p^i = m_0c^2u^i + e\phi^i. \tag{23-13}$$

Hence,

$$H_4 = \frac{1}{2}m_0c^2u^iu_i = \frac{1}{2m_0c^2}\{[p^i - e\phi^i(x^i)][p_i - e\phi_i(x^i)]\} \tag{23-14}$$

gives the explicit functional dependence of H_4 on p^i and x^i. Numerically, however, $u^i u_i = 1$, as can be easily verified by means of the components of u^i, namely, $u^i = (\mathbf{u}/c,1)/\sqrt{1 - u^2/c^2}$.

The four-Hamiltonian, Eq. (23–14), together with Eq. (23–11), gives the correct equation of motion. By direct differentiation,

$$\begin{aligned}\frac{\partial H_4}{\partial q^i} &= -\frac{e}{m_0 c^2}(p^j - e\phi^j)\frac{\partial \phi_j}{\partial x^i} \\ &= -\frac{dp_i}{ds} = -m_0 c^2 \frac{du_i}{ds} - e\frac{\partial \phi_i}{\partial x^j}u^j. \end{aligned} \tag{23–15}$$

Hence,

$$eu^j\left(\frac{\partial \phi_j}{\partial x^i} - \frac{\partial \phi_i}{\partial x^j}\right) = m_0 c^2 \frac{du_i}{ds}, \tag{23–16}$$

in agreement with Eq. (23–1).

Let us now turn to the ordinary three-dimensional representation of the Hamiltonian. We postulate

$$H(p_\alpha,q_\alpha,t) = p^4(p_\alpha,q_\alpha,x^4), \tag{23–17}$$

where p^4 is the solution of Eq. (23–14) for $H_4 = \frac{1}{2}m_0c^2$, i.e.,

$$H(p_\alpha,q_\alpha,t) = e\phi + c\sqrt{(\mathbf{p} - e\mathbf{A})^2 + (m_0c)^2}. \tag{23–18}$$

If we put

$$\begin{aligned} L &= -H + \mathbf{p}\cdot\mathbf{u}, \\ \mathbf{p} - e\mathbf{A} &= \frac{m_0\mathbf{u}}{\sqrt{1 - u^2/c^2}} = m\mathbf{u}, \end{aligned} \tag{23–19}$$

we find that

$$L(u,q_\alpha,t) = -e\phi + e(\mathbf{A}\cdot\mathbf{u}) - m_0c^2\sqrt{1 - u^2/c^2} \tag{23–20}$$

is the ordinary three-dimensional Lagrangian. Note that this Lagrangian fills the necessary relativistic requirement that the action integral,

$$\int L\,dt = -\int (u^i\phi_i + m_0c^2)\,ds, \tag{23–21}$$

shall be a scalar invariant. It can be easily shown that the three-dimensional Hamiltonian and Lagrangian, Eqs. (23–18) and (23–20) respectively, lead to the equation of motion, (23–2).

It remains to be shown that if $p^4 = H$, then H satisfies Eq. (23–10). According to Eq. (23–11), we can establish the relation between x^4 and s, i.e., between "time" and "proper time":

$$\frac{dx^4}{ds} = +\frac{\partial H_4}{\partial p^4}. \tag{23–22}$$

Hence, in general,

$$\frac{dx_\alpha}{dx^4} = -\frac{\partial H_4/\partial p_\alpha}{\partial H_4/\partial p_4}, \quad \frac{dp_\alpha}{dx^4} = +\frac{\partial H_4/\partial x_\alpha}{\partial H_4/\partial p_4}. \tag{23-23}$$

[Note the change in sign due to the presence of covariant components in Eq. (23–11).] If we set the total derivative of the invariant in Eq. (23–14) equal to zero, we obtain

$$\frac{\partial H_4}{\partial x_\alpha} + \frac{\partial H_4}{\partial p_4}\frac{\partial p_4}{\partial x_\alpha} = 0, \tag{23-24}$$

$$\frac{\partial H_4}{\partial p_\alpha} + \frac{\partial H_4}{\partial p_4}\frac{\partial p_4}{\partial p_\alpha} = 0. \tag{23-25}$$

Hence, with Eq. (23–23),

$$\frac{dx_\alpha}{dt} = \frac{\partial p_4}{\partial p_\alpha}, \tag{23-26}$$

$$\frac{dp_\alpha}{dt} = -\frac{\partial p_4}{\partial x_\alpha}. \tag{23-27}$$

Thus $p_4(x_\alpha,p_\alpha)$ satisfies the ordinary three-dimensional Hamiltonian equation, and the postulated Eq. (23–17) leads to a relativistically correct Hamiltonian and to a correspondingly correct Lagrangian.

23–3 Equations for the trajectories. Thus far we have assembled the various relativistically correct expressions for integrating the equations of motion with respect to the time. Frequently, however, there are cases where it is desirable to obtain equations for the particle orbits, without regard to the time involved in the course of the path. These can be found by integrating the equations of motion in time and then eliminating the time but, in general, it is less tedious to obtain directly differential equations for the orbits in space. This can be accomplished most easily by means of the "principle of least action."

It should be recalled that in classical mechanics there are two variational principles in general use: (1) Hamilton's principle, $\delta\int_{t_1}^{t_2} L\,dt = 0$, and (2) the principle of least action, $\delta\int_{x_{\alpha_1}}^{x_{\alpha_2}} \mathbf{p}\cdot d\mathbf{l} = 0$. These principles are quite different in their physical content. We have already made use of Hamilton's principle, in which the variation is taken between two unvaried points in space and time; the varied paths will not obey the equations of motion or the conservation laws. In the principle of least action, the end points are fixed points in space but *not* in time; however, the varied path does obey the law of conservation of energy.

Together with Eq. (23–19) the principle of least action gives

$$\delta\int_{x_\alpha}^{x_{\alpha_2}} \mathbf{p}\cdot d\mathbf{l} = \delta\int_{x_{\alpha_1}}^{x_{\alpha_2}} (m\mathbf{u} + e\mathbf{A})\cdot d\mathbf{l}$$

$$= \delta\int_{x_{\alpha_1}}^{x_{\alpha_2}} \left(\frac{m_0\mathbf{u}}{\sqrt{1 - u^2/c^2}} + e\mathbf{A}\right)\cdot d\mathbf{l} = 0. \tag{23–28}$$

Let us now make one of the three coordinates, say x, the independent variable, and let y', z' be the slopes dy/dx, dz/dx, respectively. Equation (23–19) then becomes

$$\delta\int_{x_1}^{x_2} [mu\sqrt{1 + y'^2 + z'^2} + e(A_x + A_y y' + A_z z')]\,dx = 0. \tag{23–29}$$

The Eulerian equations of this system will yield the differential equations for the orbit if the magnitude of the velocity u is known as a function of x. Since in the principle of least action the varied paths obey the law of conservation of energy, we can express u in terms of the kinetic energy T and the rest energy E_0 of the particle, i.e.,

$$\frac{u}{c}\frac{1}{\sqrt{1 - u^2/c^2}} = \sqrt{T^2 + 2TE_0}/E_0.$$

In terms of $\tau = T/E_0$ the variational equation is then given by

$$\delta\int_{\tau_1}^{\tau_2}\left[\sqrt{2\tau(1 + \tau/2)(1 + y'^2 + z'^2)} + \frac{ec}{E_0}(A_x + A_y y' + A_z z')\right]dx = 0. \tag{23–30}$$

The principal application of these considerations is for the calculation of orbits in electric and magnetic lens structures. This is a fairly specialized subject, and only solutions of general interest will be discussed here. For more detailed information the reader is referred to the electron optics literature, some of which is listed at the end of the chapter.

In electron and ion optics one usually deals with orbits in the proximity of an axis. Let us first consider the case of radial symmetry about such an axis, i.e., the only component of $\mathbf{A}$ is A_θ, which is independent of θ. If, in cylindrical coordinates, we make z the independent variable, with $r' = dr/dz$ and $\theta' = d\theta/dz$, Eq. (23–30) becomes

$$\delta\int_{z_1}^{z_2}\left[\sqrt{2\tau(1 + \tau/2)}\,(r'^2 + r^2\theta'^2 + 1)^{1/2} + \frac{ec}{E_0}A_\theta r\theta'\right]dz = 0.$$

Eulerian equations for $r(z)$ and $\theta(z)$ can then be generated in the usual way. A first integral of the equation in $\theta(z)$ can be found immediately and leads to the law of "conservation of canonical angular momentum" in the form

$$\frac{r^2\theta'\sqrt{2\tau(1+\tau/2)}}{(r'^2+r^2\theta'^2+1)^{1/2}}+\frac{ec}{E_0}A_\theta r = C. \tag{23–31}$$

It can be shown that Eq. (23–31) is equivalent to

$$rp_\theta + eA_\theta r = C', \tag{23–32}$$

where $p_\theta = r\dot{\theta}m_0/\sqrt{1-u^2/c^2}$. Hence a particle emitted in a magnetic field will attain mechanical angular momentum when leaving the field; a particle passing through a magnetic field will experience a change in mechanical angular momentum, but this will be restored to its initial value on leaving the field.

Equation (23–31) can be substituted into the Eulerian differential equation for $r(z)$ in order to eliminate dependence on θ'. The resulting equation is algebraically complicated, but it can be stated simply by means of a change in parameters. Let

$$\begin{aligned} \xi &= 2\tau(1+\tau/2), \\ \eta &= \frac{1}{\xi^{1/2}}\left(\frac{C}{r}-\frac{ec}{E_0}A_\theta\right), \\ \lambda &= \xi(1-\eta^2), \end{aligned} \tag{23–33}$$

where C is the constant of Eq. (23–31). Then, without any approximation, the Eulerian equation becomes

$$r'' = \frac{1+r'^2}{2\lambda}\left[(1+r'^2)\frac{\partial\lambda}{\partial r} - r'\frac{\partial\lambda}{\partial z}\right]. \tag{23–34}$$

This equation is more general than is usually necessary. In most cases, λ has a simple form; e.g., for a purely electrostatic system and low velocity motion λ is just the kinetic energy and hence the electrostatic potential.

Equation (23–34) reduces to a much simpler form if one considers "paraxial" motion, i.e., motion in which the inclination of rays to the axis never becomes large. In this case, two approximations are possible: we can ignore r'^2, and we can relate $\partial/\partial r$ to $\partial/\partial z$ by the field equations. Specifically, we have, near the axis,

$$\frac{\partial\phi}{\partial r} = -\frac{r}{2}\frac{\partial^2\phi}{\partial z^2}, \tag{23–35}$$

and

$$A_\theta = \frac{Br}{2}, \tag{23–36}$$

correct to order r^2, and we need consider the values of ϕ and B only on the axis. With these approximations, Eq. (23–34) becomes

$$r'' + r'\frac{\tau'}{2\tau}\left(\frac{1+\tau}{1+\tau/2}\right) + \frac{r}{4}\frac{\tau''}{\tau}\left(\frac{1+\tau}{1+\tau/2}\right) + \frac{r}{8}\frac{(eB/m_0c)^2}{\tau(1+\tau/2)}$$

$$- \frac{C^2}{2\tau(1+\tau/2)r^3} = 0. \qquad (23\text{–}37)$$

If we put $\beta = u/c$ and $\gamma = 1/\sqrt{1 - u^2/c^2}$, Eq. (23–37) takes the simpler form

$$r'' + \frac{r'\gamma'}{\beta^2\gamma} + \frac{r\gamma''}{2\beta^2\gamma} + \frac{r}{4}\left(\frac{Be}{m_0c\beta\gamma}\right)^2 - \frac{C^2}{\beta^2\gamma^2r^3} = 0. \qquad (23\text{–}38)$$

The physical content of Eq. (23–38) is clear. The first three terms correspond to an electric lens system with orbits of zero angular momentum. The last term adds a "centrifugal potential" or an effective "repulsive core" to the system in case the canonical angular momentum is different from zero. The fourth term represents the lens action of a magnetic field. We shall consider the effects separately.

23–4 Applications.

1. *Electrostatic lens, no "eccentric" rays, $B = 0 = C$.* In this case, Eq. (23–38) becomes

$$r'' + \frac{r'\gamma'}{\beta^2\gamma} + \frac{r\gamma''}{2\beta^2\gamma} = 0, \qquad (23\text{–}39)$$

or

$$(\beta\gamma r')' + \frac{r\gamma''}{2\beta} = 0. \qquad (23\text{–}40)$$

Qualitatively, the action of an electrostatic lens depends on the change in velocity of a particle traversing the system. In Fig. 23–2 it is assumed that a positive charge travels near the axis from left to right. In an accelerating lens the particle is faster and hence spends less time in the

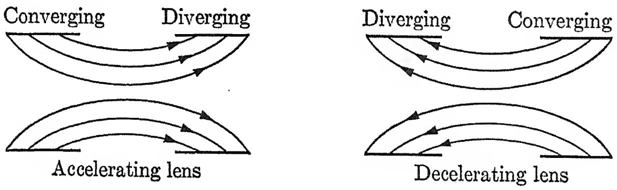


FIG. 23–2 Accelerating and decelerating electrostatic lenses, both of net convergence.

diverging part of the field; conversely, if the lens produces deceleration the particle is slower in the now converging second part of the field. In either case, therefore, the lens is converging. This can be seen quantitatively if we apply the transformation

$$R = rT^{1/4} \tag{23-41}$$

to Eq. (23-39) in the nonrelativistic approximation. We obtain

$$R'' + \frac{3}{16}\left(\frac{T'}{T}\right)^2 R = 0, \tag{23-42}$$

which demonstrates the convergence property of the lens. By integrating Eq. (23-42) through a region of variable potential, we can obtain formulas for the focal properties of electrostatic lenses.

The velocity of the particle, and hence the (nonrelativistic) square root of the electrostatic potential, is analogous to the refractive index, n, in optics. This can be understood immediately if we note the correspondence between Fermat's principle in optics, $\delta\int_{x_1}^{x_2} n\,dl = 0$, and the nonrelativistic principle of least action, $\delta\int_{x_1}^{x_2} \sqrt{T}\,dl = 0$. Note also that electrostatic lenses are "second order" devices in the sense that the focusing action depends quadratically on T' and hence quadratically on the axial field strength; physically, this is due to the fact that the lens action depends on *both* the change in velocity caused by the field *and* the radial component of the field.

For simplicity, we have considered only the nonrelativistic case. The same arguments apply to the general case, except that the transformation leading to Eq. (23-42) is somewhat more complicated than Eq. (23-41).

2. *Magnetostatic lens, no "eccentric" rays,* $C = 0$. If there are no electrostatic fields to produce changes in energy, Eq. (23-38) becomes simply

$$r'' + \frac{r}{4}\left(\frac{B}{B\rho}\right)^2 = 0, \tag{23-43}$$

where $B\rho = cm_0\beta\gamma/e$ is the "magnetic rigidity" of the particle, i.e., the product of the magnetic field and the radius of curvature. No approximation involving small velocities needs to be made in this case. It is obvious that a magnetic lens, like an electrostatic lens, is always a converging device. Equation (23-43) can be readily integrated to give the focal length.

The lens action is also of second order in this case, i.e., it is quadratic in the field. Let us try to understand this physically. Consider a magnetic lens, say in the form of a solenoid, and let a particle enter with zero canon-

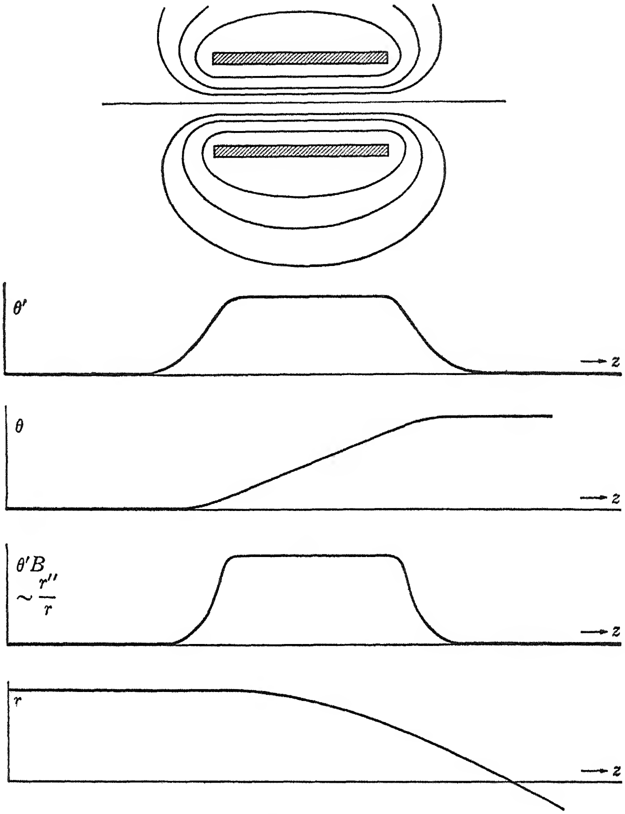


FIG. 23-3 Effect of a magnetic lens.

ical angular momentum (Fig. 23-3). While entering the fringing field, the particle will acquire a mechanical angular momentum

$$-p_\theta = eA_\theta \simeq \frac{eBr}{2}, \tag{23-44}$$

according to Eq. (23-32); this is also qualitatively evident by inspection of the Lorentz force on the charge in this part of the field. This angular velocity then interacts with the longitudinal component of **B** to produce a radially inward force. Thus **B** has acted twice, and the net effect is quadratic. The reader can easily convince himself that the direction of the net radial force is always opposite to the direction of the radius vector.

3. *General case.* Equation (23-38) can be written as

$$\frac{d}{dz}(\beta\gamma r') + r\left\{\frac{\gamma''}{2\beta} + \frac{1}{4}\left(\frac{eB}{m_0c}\right)^2\frac{1}{\beta\gamma} - \frac{C^2}{\beta\gamma r^4}\right\} = 0, \tag{23-45}$$

which, except for the last term, is the equation of motion corresponding to a harmonic oscillator of variable mass and variable force constant. The last term corresponds to an r^{-2} repulsive core in such a potential; finite canonical angular momentum makes it impossible for the particle to cross the axis.

In general, the solution of Eq. (23–45) requires numerical or graphical integration. Note that β and γ are assumed to be given functions of z through the value of the electrostatic potential on the axis. If $C = 0$ and the parameters vary slowly with respect to the period of radial oscillations, the equation can be integrated for extended systems by means of the so-called "adiabatic theorem." In that case the equation has the form

$$(Fr')' + Gr = 0. \tag{23–46}$$

If F and G are functions of z fulfilling the conditions that

$$F'/F \ll \sqrt{G/F}, \quad G'/G \ll \sqrt{G/F},$$

solutions of Eq. (23–46) are represented by

$$r = \frac{1}{\sqrt[4]{FG}}\, e^{\pm i \int^{z} \sqrt{G/F}\, dz}, \tag{23–47}$$

and the exponent may be integrated analytically or numerically, depending on the nature of the fields.

4. *First-order lenses.* It can be easily shown that it is impossible to build a first-order lens, either electric or magnetic, which has cylindrical symmetry. In an electrostatic field of cylindrical symmetry the mean value of E_r will always vanish in charge-free regions because of the divergence condition. Similarly, the mean value of the azimuthal component B_θ of the magnetic field will vanish in a current-free region, since in this case **B** is irrotational. Hence, no net radial impulse of first order in **E** or **B** can be given to a particle in a field of cylindrical symmetry.

First-order focusing effects can be produced by asymmetrical lenses. The prime example of such a lens is the electric or magnetic quadrupole, which is essentially a "fully astigmatic" lens. It can be seen from Fig. 23–4 that such a structure, if acting as a thin lens, will have equal but opposite focal strengths in two mutually perpendicular planes; hence the

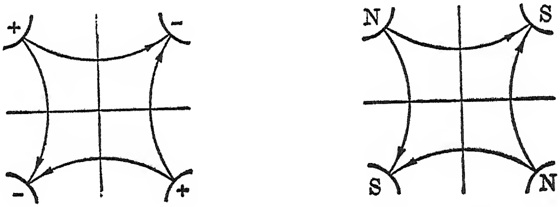


FIG. 23–4 Electric and magnetic quadrupole lenses.

theorem that the mean radial momentum imparted to a particle (averaged over azimuthal angles) shall vanish is still not violated. The combination of a diverging lens and a converging lens of equal strength separated in space is converging; hence lens combinations made up of alternating elements of quadrupoles are converging. This is the basis of the "strong-focusing" principle in electron and ion optics.

5. *Solutions simplified by change of Lorentz frame.* Frequently, integration of the equations of motion may be simplified by choosing a Lorentz frame other than the frame of physical interest. Most notable are the cases where the combined effect of electric and magnetic fields can be reduced to the effect of a purely electric or a purely magnetic field. The possibilities for such transformations can be ascertained by examination of the two invariants, scalar and pseudoscalar,

$$E^2 - c^2B^2 = I_1, \tag{23-48}$$

$$\mathbf{E} \cdot \mathbf{B} = I_2. \tag{23-49}$$

We have seen in Chapter 17 that if $I_1 = 0 = I_2$, we have the case of a plane wave which is invariant in all its properties. If $I_2 = 0$ but $I_1 \neq 0$, it is possible to transform away either **B** or **E**. If $E < cB$, a frame can be found for which **E** vanishes. If **B** is uniform, the orbits are then circles (or helixes) and thus become cycloids in the laboratory frame. If $cB < E$, a transformation may be effected to a frame in which **B** vanishes, and the orbits are catenaries if **E** is uniform. A component of **E** parallel to **B** ($I_2 \neq 0$) cannot be transformed away, but if **E** is parallel to **B**, motion in the plane defined by the fields and the initial velocity depends only on **E** and can be integrated first; a solution for the motion at right angles to this plane then follows.

SUGGESTED REFERENCES

H. GOLDSTEIN, *Classical Mechanics.* There are many adequate treatments of Lagrangian and Hamiltonian particle mechanics, but the reader may find this modern work most useful.

H. C. CORBEN AND P. M. STEHLE, *Classical Mechanics.* Another modern treatment of classical mechanics. There is very little relativistic mechanics in this work, but the distinction between covariant and contravariant components of vectors and tensors is maintained so that the formulation is immediately useful in relativistic theory.

L. LANDAU AND E. LIFSHITZ, *The Classical Theory of Fields.* Chapter 3 is devoted to charges in fields, and covers much of the material we have included here.

O. KLEMPERER, *Electron Optics*, 2nd edition. In addition to a discussion of the principles and many applications, this expanded version of an early work also contains an excellent bibliography.

V. K. ZWORYKIN, G. A. MORTON, E. G. RAMBERG, J. HILLIER, AND A. W. VANCE, *Electron Optics and the Electron Microscope.* A fairly complete treatment of general interest as of 1945.

Two bibliographies of electron microscopy appeared in 1950, one edited by V. E. Coslett, the other by Claire Marten *et al.*

EXERCISES

1. Verify that the Lagrangian

$$L = -m_0c^2\sqrt{1 - u^2/c^2} - e\phi + e\mathbf{A}\cdot\mathbf{u}$$

yields the correct equation of motion.

2. Show that the orbit of a particle of charge e, rest mass m_0, in a uniform electrostatic field $\mathbf{E}$ along the x-axis, is

$$x = \frac{U_0}{eE}\cosh\frac{eEy}{p_0c},$$

where $U_0 = \sqrt{(m_0c^2)^2 + c^2p_0^2}$ and p_0 is the initial momentum in the y-direction. The momentum in the x-direction at $y = 0$ has been taken equal to zero.

3. Consider an accelerating column for charged particles such as might be installed in an electrostatic generator. As a result, the energy of the particle is given approximately by

$$e\phi = k\left(z + \frac{g\alpha}{2\pi}\sin\frac{2\pi z}{g}\right),$$

$$e\phi' = k\left(1 + \alpha\cos\frac{2\pi z}{g}\right),$$

where z is the axial distance and k is the mean accelerating field. Since

$$1 + \frac{\alpha^2}{2} = \frac{\overline{E^2}}{\bar{E}^2},$$

α is a measure of the fluctuation of the axial field. Assume that g, the wavelength of these variations, is small compared with the focal length of each lens of thickness g. Hence, ϕ'^2 may be represented by its mean square and the fluctuation term in ϕ may be neglected. What is the minimum value of α such that the orbits will be oscillatory?

4. An electrostatic "strong-focusing" lens produces a field corresponding to the complex potential

$$W = \frac{E_0z^2}{2a},$$

where E_0 is the electric field at radius a. How could such a field be realized? If the length of the lens is L, calculate the focal length in the x- and y-planes. Do *not* assume the lens to be thin.

5. Consider two coaxial conducting circular cylinders of equal radii charged to potentials $-V_0/2$, $+V_0/2$, respectively. Let the separation between the cylinders be small compared with the radius a of the cylinders. Calculate the potential within the space bounded by the cylinders.

6. In the preceding problem, let $V_0 = A\cos\omega t$, where $\omega \ll c/a$, so that the static solution is valid. Let a particle of charge e cross the field along a trajectory parallel to the axis at a distance b from it. What is the energy gain?

7. Consider a particle of energy $U_0 \gg eV_0$ passing through the system of problem 5. What is the focal length?

8. Prove Eq. (23–47) as a solution of Eq. (23–46).

CHAPTER 24

HAMILTONIAN FORMULATION OF MAXWELL'S EQUATIONS

In Chapter 23 we found the covariant equations of motion for a point particle in an external field, using the Lagrangian and Hamiltonian formulations. Let us now formulate the equations of motion of the field, i.e., Maxwell's equations, in terms of Lagrangian and Hamiltonian functions. No new physical information concerning classical electrodynamics can be arrived at in this way, but the transition to quantum electrodynamics is accomplished through canonical formulation of the field. Moreover, the methods facilitate the investigation of various types of fields other than the electromagnetic.

In extending the Hamiltonian formulation to a field, we are faced with a new consideration, namely, the fact that the number of degrees of freedom, in the mechanical sense, is infinite. We shall begin, therefore, by studying the transition of a mechanical system of N degrees of freedom to a system in which N becomes infinite.

24–1 Transition to a one-dimensional continuous system. Let us consider a set of N point particles of equal mass m, connected by springs of equal length a and force constant k. (See Fig. 24–1.) Let η_i be the displacement from equilibrium of the ith mass. The solution of this problem rests on finding a suitable Lagrangian L such that the equation

$$\delta \int L(\eta_i, \dot{\eta}_i, t)\, dt = 0 \tag{24–1}$$

represents the correct equation of motion. In classical mechanics, we know that $L = T - V$, where T and V are the kinetic and potential energies respectively. By the geometry of the problem at hand, we thus have (neglecting end contributions):

$$L = \tfrac{1}{2} \sum_{i=1}^{N} [m\dot{\eta}_i^2 - k(\eta_{i+1} - \eta_i)^2]. \tag{24–2}$$

Equation (24–2) can be written in the form

$$L = \sum_{i=1}^{N} a\mathcal{L}_i, \tag{24–3}$$

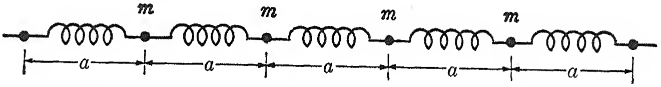


FIG. 24–1 A system of mass points connected by springs.

where

$$\mathfrak{L}_i = \frac{1}{2}\left[\frac{m}{a}\dot{\eta}_i^2 - ka\left(\frac{\eta_{i+1}-\eta_i}{a}\right)^2\right] \tag{24–4}$$

is a quantity which we might call the "linear Lagrangian density." Equation (24–1), with the Lagrangian of Eq. (24–2), gives the usual equation of motion for a set of coupled oscillators; orthogonalization of Eq. (24–2) yields the set of normal modes.

Our principal interest here is to allow the number of degrees of freedom to tend to infinity; this we shall do by letting

$$\begin{aligned} a &\to dx, \\ m/a &\to \mu, \\ ka &\to Y, \\ \frac{\eta_{i+1}-\eta_i}{a} &\to \frac{\partial\eta}{\partial x}, \end{aligned} \tag{24–5}$$

where μ is the linear mass density and Y is Young's modulus. By means of this transition, we have replaced the discrete index i by the continuous variable x. The Lagrangian then becomes

$$L = \frac{1}{2}\int\left[\mu\dot{\eta}^2 - Y\left(\frac{\partial\eta}{\partial x}\right)^2\right]dx = \int \mathfrak{L}\,dx, \tag{24–6}$$

with

$$\mathfrak{L}\left(\eta,\dot{\eta},\frac{\partial\eta}{\partial x},t\right) = \frac{1}{2}\left[\mu\dot{\eta}^2 - Y\left(\frac{\partial\eta}{\partial x}\right)^2\right]. \tag{24–7}$$

The Lagrangian density thus acquires an explicit dependence on the spatial derivatives of the "field coordinate" η.

Let us now derive the Eulerian equation corresponding to the variational principle of Eq. (24–1), which has become

$$\delta\int L\,dt = \delta\iint \mathfrak{L}\,dx\,dt = 0. \tag{24–8}$$

On substituting Eq. (24–7) and integrating by parts in the usual way, we obtain (ignoring explicit time dependence)

$$\begin{aligned} \delta\int L\,dt &= \iint\left\{\frac{\partial\mathfrak{L}}{\partial\eta}\,\delta\eta + \frac{\partial\mathfrak{L}}{\partial(\partial\eta/\partial x)}\,\delta\left(\frac{\partial\eta}{\partial x}\right) + \frac{\partial\mathfrak{L}}{\partial(\partial\eta/\partial t)}\,\delta\left(\frac{\partial\eta}{\partial t}\right)\right\}dx\,dt \\ &= \iint\left\{\frac{\partial\mathfrak{L}}{\partial\eta} - \frac{\partial}{\partial x}\left[\frac{\partial\mathfrak{L}}{\partial(\partial\eta/\partial x)}\right] - \frac{\partial}{\partial t}\left[\frac{\partial\mathfrak{L}}{\partial(\partial\eta/\partial t)}\right]\right\}\delta\eta\,dx\,dt = 0. \end{aligned} \tag{24–9}$$

The integrated part has been made to vanish by the condition that $\delta\eta = 0$ at the end points of integration over t. Since $\delta\eta$ is an arbitrary function of x, we obtain the partial differential equation

$$\frac{\partial \mathfrak{L}}{\partial \eta} - \frac{\partial}{\partial x}\left[\frac{\partial \mathfrak{L}}{\partial(\partial\eta/\partial x)}\right] - \frac{\partial}{\partial t}\left[\frac{\partial \mathfrak{L}}{\partial(\partial\eta/\partial t)}\right] = 0. \tag{24-10}$$

Equation (24–10) is often written in the form

$$\frac{\delta \mathfrak{L}}{\delta \eta} - \frac{\partial}{\partial t}\left[\frac{\partial \mathfrak{L}}{\partial(\partial\eta/\partial t)}\right] = 0, \tag{24-11}$$

where

$$\frac{\delta \mathfrak{L}}{\delta \eta} = \frac{\partial \mathfrak{L}}{\partial \eta} - \frac{\partial}{\partial x}\left[\frac{\partial \mathfrak{L}}{\partial(\partial\eta/\partial x)}\right] \tag{24-12}$$

is called the "variational" or "functional" derivative. Note that in going to the limit of a continuous variable we have replaced a system of N ordinary differential Lagrangian equations by a partial differential equation.

With the substitution of the Lagrangian density, Eq. (24–7), Eq. (24–11) yields immediately the wave equation,

$$\mu \frac{\partial^2 \eta}{\partial t^2} - Y \frac{\partial^2 \eta}{\partial x^2} = 0, \tag{24-13}$$

corresponding to compressional waves traveling with velocity $\sqrt{Y/\mu}$.

24–2 Generalization to a three-dimensional continuum. These considerations may be generalized to a three-dimensional field η, where η may be any covariant parameter. The variational principle corresponding to Eq. (24–1) is given by

$$\delta \iint \mathfrak{L}\, dv\, dt = 0, \tag{24-14}$$

or

$$\delta \iint \mathfrak{L}\left(\eta, \frac{\partial \eta}{\partial x^i}\right) d^4x = 0. \tag{24-15}$$

This formulation is evidently covariant if the Lagrangian density $\mathfrak{L}$ is a scalar. Let us again vary the functional dependence of $\mathfrak{L}$ on η and $\partial\eta/\partial x^i$, but consider the x^i as fixed independent coordinates. Partial integration gives

$$\int \left\{\frac{\partial \mathfrak{L}}{\partial \eta} - \frac{\partial}{\partial x^i}\left[\frac{\partial \mathfrak{L}}{\partial(\partial\eta/\partial x^i)}\right]\right\} d^4x\, \delta\eta = 0, \tag{24-16}$$

leading to the covariant Lagrangian equations

$$\frac{\partial \mathfrak{L}}{\partial \eta} = \frac{\partial}{\partial x^i}\left[\frac{\partial \mathfrak{L}}{\partial(\partial\eta/\partial x^i)}\right]. \tag{24–17}$$

To exhibit the time dependence explicitly, we may write these equations in the form

$$\frac{\delta \mathfrak{L}}{\delta \eta} = \frac{\partial}{\partial t}\left(\frac{\partial \mathfrak{L}}{\partial \dot{\eta}}\right), \tag{24–18}$$

where

$$\frac{\delta \mathfrak{L}}{\delta \eta} = \frac{\partial \mathfrak{L}}{\partial \eta} - \frac{\partial}{\partial x^\alpha}\left[\frac{\partial \mathfrak{L}}{\partial(\partial\eta/\partial x^\alpha)}\right], \tag{24–19}$$

with $\dot{\eta} = \partial\eta/\partial t$ and with x^α representing the three space variables. Our program is to find an $\mathfrak{L}$ such that Eq. (24–17) will lead, in the electromagnetic case, to Maxwell's equations. Clearly, this discussion will apply to any field theory.

It has been possible to state the action principle and write a Lagrangian partial differential equation in obviously covariant form. To introduce a Hamiltonian, we must single out the time among the x^i as has been done in Eqs. (24–18) and (24–19). We can then define a "momentum density" conjugate to η,

$$\pi(x^\alpha,t) = \frac{\partial \mathfrak{L}}{\partial \dot{\eta}}, \tag{24–20}$$

and a Hamiltonian density,

$$\mathfrak{H}\left(\eta, \frac{\partial\eta}{\partial x_\alpha}; \pi, \frac{\partial\pi}{\partial x_\alpha}; t\right) = \pi\dot{\eta} - \mathfrak{L}. \tag{24–21}$$

The Hamiltonian equations follow in the usual way. Consider an increment ∂H of the total Hamiltonian $H = \int \mathfrak{H}\, dv$:

$$\partial H = \int\left[\dot{\eta}\,\partial\pi + \pi\,\partial\dot{\eta} - \frac{\partial\mathfrak{L}}{\partial(\partial\eta/\partial x^i)}\,\partial\left(\frac{\partial\eta}{\partial x^i}\right) - \frac{\partial\mathfrak{L}}{\partial\eta}\,\partial\eta\right] dv.$$

With Eq. (24–17) and the definition of π, this becomes, on partial integration,

$$\partial H = \int (\dot{\eta}\,\partial\pi - \dot{\pi}\,\partial\eta)\, dv. \tag{24–22}$$

Since

$$H = \int \mathfrak{H}\left(\pi, \frac{\partial\pi}{\partial x^\alpha}, \eta, \frac{\partial\eta}{\partial x^\alpha}\right) dv, \tag{24–23}$$

we may also write

$$\partial H = \int \left[\frac{\partial \mathcal{H}}{\partial \pi} \partial \pi + \frac{\partial \mathcal{H}}{\partial(\partial \pi/\partial x^{\alpha})} \partial \left(\frac{\partial \pi}{\partial x^{\alpha}} \right) + \frac{\partial \mathcal{H}}{\partial \eta} \partial \eta + \frac{\partial \mathcal{H}}{\partial(\partial \eta/\partial x^{\alpha})} \partial \left(\frac{\partial \eta}{\partial x^{\alpha}} \right) \right] dv. \tag{24–24}$$

If we integrate Eq. (24–24) by parts and identify the result with Eq. (24–22), we obtain, in the notation of Eq. (24–19),

$$\dot{\eta} = \frac{\delta \mathcal{H}}{\delta \pi}, \quad \dot{\pi} = - \frac{\delta \mathcal{H}}{\delta \eta} \tag{24–25}$$

as the new form of Hamilton's equations.

Hamilton's equations lead to the usual meaning of the time rate of change in terms of Poisson brackets. If Λ is the density of a physical variable L, i.e., if $L = \int \Lambda \, dv$, then by means of the process used for obtaining Eq. (24–25), we find

$$\begin{aligned} \frac{dL}{dt} &= \int \left(\frac{\delta \Lambda}{\delta \eta} \dot{\eta} + \frac{\delta \Lambda}{\delta \pi} \dot{\pi} \right) dv \\ &= \int \left(\frac{\delta \Lambda}{\delta \eta} \frac{\delta \mathcal{H}}{\delta \pi} - \frac{\delta \Lambda}{\delta \pi} \frac{\delta \mathcal{H}}{\delta \eta} \right) dv \\ &= \int [\Lambda, \mathcal{H}] \, dv, \end{aligned} \tag{24–26}$$

which is analogous to the usual Poisson bracket. These expressions lead to convenient starting points for quantization.

The discussion above leads to definite field equations if a Lagrangian density $\mathcal{L}$ is given. In order that the field equations be linear, the Lagrangian must not contain higher powers than the second of either η or $\partial \eta / \partial x^i$. As the simplest example, we might consider

$$\mathcal{L} = \frac{1}{2} \left(\frac{\partial \eta}{\partial x^i} \frac{\partial \eta}{\partial x_i} - \mu^2 \eta^2 \right), \tag{24–27}$$

which leads, by Eq. (24–11), to the field equations

$$(\Box - \mu^2)\eta = 0. \tag{24–28}$$

The corresponding momentum density is given by

$$\pi = \frac{1}{c^2} \frac{\partial \eta}{\partial t}, \tag{24–29}$$

and we are led to a positive definite Hamiltonian,

$$\mathcal{H} = \tfrac{1}{2}[c^2\pi^2 + (\nabla\eta)^2 + \mu^2\eta^2]. \tag{24-30}$$

This is the scalar meson field of Yukawa, of which the point solution is

$$\eta = e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\,\frac{e^{-\mu r}}{r}. \tag{24-31}$$

24-3 The electromagnetic field. The electromagnetic field demands a more complicated formulation. Presumably, we are now dealing with a vector field, i.e.,

$$\eta \to \phi^i. \tag{24-32}$$

The total Lagrangian will have three parts: (1) a mass term for the motion of material particles; (2) an interaction term relating particle to field; and (3) a field term corresponding to the field equation.

Terms (1) and (2) can be written down from our former considerations. If we take

$$\mathcal{L}_1 = \tfrac{1}{2}g_0c^2u_iu^i, \tag{24-33}$$

where g_0 is the proper mass density, and

$$\mathcal{L}_2 = j^i\phi_i, \tag{24-34}$$

we know that the correct motions of point particles in an external field will result. In writing down L_3 we may be guided by a classical analogy: let us choose L_3 such that $\mathcal{L}$ is an invariant, and also is analogous to the classical difference between potential and kinetic energy. In an electromagnetic oscillation, energy oscillates between electric and magnetic energy, just as in a mechanical oscillation energy oscillates between kinetic and potential energy. The only scalar which is quadratic in the first derivatives of the four-potential is proportional to this difference, namely, $F^{ij}F_{ij} = 2(c^2B^2 - E^2)$. (The quantity $F^{ij}G_{ij}$, where G_{ij} is the dual of F_{ij}, is excluded because it would transform differently from the remaining terms of the Lagrangian, as we have seen in Section 17-2.) It is, of course, not excluded that a Lagrangian density might also include terms quadratic in the four-potentials themselves, such as that found in Eq. (24-27), but it is just this term that introduces an exponential dependence on r into the point solution, Eq. (24-31). Thus, while a Lagrangian density analogous to Eq. (24-27) yields a possible field theory from the point of view of the transformation character of the terms, it does not correspond to the experimental properties of the electromagnetic field. When translated into quantum mechanics, a theory in which the Lagrangian density contains terms quadratic in the potentials as well as their derivatives predicts photons of finite rest mass, instead of the zero rest mass photons of electromagnetic theory.

Thus, for the electromagnetic field, we are led to take

$$\mathcal{L}_3 = \frac{\epsilon_0(E^2 - c^2B^2)}{2} = -\frac{\epsilon_0}{4}F^{ij}F_{ij}. \tag{24-35}$$

Note that the field equations,

$$\nabla \cdot \mathbf{B} = 0, \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t},$$

are already implied in the connecting equations

$$F^{ij} = \frac{\partial \phi^j}{\partial x_i} - \frac{\partial \phi^i}{\partial x_j} \tag{17-30}$$

corresponding to

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}, \quad \mathbf{E} = \nabla \times \mathbf{A}.$$

With Eq. (24–35), the total Lagrangian density is given by

$$\mathcal{L} = -\frac{\epsilon_0}{4}F^{ij}F_{ij} + eu^i\phi_i + \frac{1}{2}g_0c^2u^iu_i. \tag{24-36}$$

A further reason for this choice of the Lagrangian is the connection between Lagrangian density and the energy momentum tensor, Eq. (22–24). If we consider the Hamiltonian as retaining its meaning as an energy density, we can put

$$\frac{\partial \mathcal{H}}{\partial t} + \nabla \cdot \mathbf{N} = 0. \tag{24-37}$$

Now

$$\begin{aligned}\frac{\partial \mathcal{H}}{\partial t} &= \frac{\partial}{\partial t}\left(\dot{\eta}\frac{\partial \mathcal{L}}{\partial \dot{\eta}} - \mathcal{L}\right) = \dot{\eta}\frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\eta}}\right) - \dot{\eta}\frac{\partial \mathcal{L}}{\partial \eta} - \frac{\partial \mathcal{L}}{\partial(\partial\eta/\partial x^\alpha)}\frac{\partial^2\eta}{\partial x^\alpha \partial t} \\ &= -\frac{\partial}{\partial x^\alpha}\left[\dot{\eta}\frac{\partial \mathcal{L}}{\partial(\partial\eta/\partial x^\alpha)}\right].\end{aligned} \tag{24-38}$$

Hence, from Eq. (24–37),

$$N^\alpha = cT^{4\alpha} = \dot{\eta}\frac{\partial \mathcal{L}}{\partial(\partial\eta/\partial x^\alpha)}. \tag{24-39}$$

The covariant generalization of Eqs. (24–39) and (24–21) is

$$T^i_j = \frac{\partial \eta}{\partial x^j}\frac{\partial \mathcal{L}}{\partial(\partial\eta/\partial x^i)} - \mathcal{L}\delta^i_j. \tag{24-40}$$

Thus the choice of Lagrangian fulfills both Eq. (24–39) and $T_4^4 = \mathcal{H}$, as required.

It can be easily shown that the Lagrangian $\mathcal{L}_2 + \mathcal{L}_3$ leads to Maxwell's equations. Note that this choice of $\mathcal{L}$ is not unique. Since the equations of motion will depend on the fields F^{ij}, which in turn do not depend on the quantity $\partial\phi^i/\partial x^i$, any function of $\partial\phi^i/\partial x^i$ may be added to $\mathcal{L}$. It is customary but not mandatory to use the Lorentz condition $\partial\phi^i/\partial x^i = 0$. From Eq. (24–17) we have, using $\mathcal{L} = \mathcal{L}_2 + \mathcal{L}_3$ and $\partial\mathcal{L}/\partial\phi_j = j^j$,

$$\frac{\partial}{\partial x^i}\,\frac{\partial\mathcal{L}}{\partial(\partial\phi_j/\partial x^i)} = \frac{\epsilon_0}{4}\,\frac{\partial}{\partial x^i}\left\{\frac{\partial}{\partial(\partial\phi_j/\partial x^i)}\left[\left(\frac{\partial\phi^l}{\partial x_k} - \frac{\partial\phi^k}{\partial x_l}\right)\left(\frac{\partial\phi_l}{\partial x^k} - \frac{\partial\phi_k}{\partial x^l}\right)\right]\right\}$$

$$= \epsilon_0\,\frac{\partial F^{ij}}{\partial x^i}, \tag{24–41}$$

and thus

$$\frac{\partial F^{ij}}{\partial x^i} = \frac{j^j}{\epsilon_0}, \tag{17–34}$$

which are Maxwell's source equations.

This calculation can, of course, also be carried out using the three-dimensional form,

$$\mathcal{L} = -\tfrac{1}{2}\epsilon_0(c^2B^2 - E^2) - \mathbf{A}\cdot\mathbf{j} + \rho\phi, \tag{24–42}$$

and Eq. (24–19). In the language of three dimensions, we may take

$$\eta^\alpha = A^\alpha, \tag{24–43}$$

$$\pi^\alpha = -\epsilon_0 E^\alpha, \tag{24–44}$$

i.e., **A** and **E** are canonically conjugate. In terms of these variables, the Hamiltonian density of the field in the absence of sources is given by

$$\mathcal{H} = \epsilon_0\left(\frac{\partial\mathbf{A}}{\partial t}\right)^2 + \epsilon_0\left(\nabla\phi\cdot\frac{\partial\mathbf{A}}{\partial t}\right) + \frac{(\nabla\times\mathbf{A})^2}{2\mu_0} - \frac{\epsilon_0}{2}\left(\frac{\partial\mathbf{A}}{\partial t}\right)^2$$

$$- \epsilon_0\left(\nabla\phi\cdot\frac{\partial\mathbf{A}}{\partial t}\right) - \frac{\epsilon_0}{2}(\nabla\phi)^2$$

$$= \frac{\pi^2}{2\epsilon_0} + \frac{(\nabla\times\boldsymbol{\eta})^2}{2\mu_0} - (\boldsymbol{\pi}\cdot\nabla\phi). \tag{24–45}$$

The last term can be made zero by a particular choice of gauge; in any case, since $\nabla\cdot\mathbf{E} = 0$, the volume integral of the last term vanishes, and thus does not contribute to the energy.

24–4 Periodic solutions in a box. Plane wave representation. An alternative way of treating Maxwell's equations in Hamiltonian form is to make the number of degrees of freedom finite by confining the field to a box of dimensions L. We may take the boundary conditions such that the field functions shall be periodic with period L in the three dimensions. The method is of great historical significance in the derivation of the Rayleigh-Jeans law (see Section 21–9), and leads to a representation of the field which is readily quantized.

Let us take as the Hamiltonian

$$H = \int \left(\frac{\epsilon_0 E^2 + \mu_0 H^2}{2} \right) dv = \int \left[\frac{\pi^2}{2\epsilon_0} + \frac{(\nabla \times \mathbf{A})^2}{2\mu_0} \right] dv. \tag{24–46}$$

Since the solutions are periodic, they can be expanded in space by a Fourier series. That is to say, if we put

$$\mathbf{u}_{k\lambda} = \frac{1}{L^{3/2}} \boldsymbol{\varepsilon}_{k\lambda} e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{24–47}$$

then in terms of these functions any vector function can be expanded as a sum over $\mathbf{k}$ and over the unit vectors, $\boldsymbol{\varepsilon}_{k\lambda}$. The boundary conditions restrict $\mathbf{k}$ to the values $\mathbf{k} = \frac{2\pi}{L}(l\hat{\mathbf{x}} + m\hat{\mathbf{y}} + n\hat{\mathbf{z}})$, with l, m, n representing all positive and negative integers, and the $\boldsymbol{\varepsilon}_{k\lambda}$ $(\lambda = 1,2,3)$ permit an arbitrary choice of polarization. For example, we can write

$$\mathbf{A}(x^\alpha,t) = \sqrt{\mu_0} \sum_{k,\lambda} \mathbf{u}_{k\lambda}(r) q_{k\lambda}(t), \tag{24–48}$$

$$\boldsymbol{\pi}(x^\alpha,t) = \sqrt{\epsilon_0} \sum_{k,\lambda} \mathbf{u}_{k\lambda}(r) p_{k\lambda}(t). \tag{24–49}$$

The functions of Eq. (24–47) satisfy the auxiliary relations,

$$\begin{aligned} \nabla \cdot \mathbf{u}_{k\lambda} &= \mathbf{k} \cdot \mathbf{u}_{k\lambda}, \\ \nabla \times \mathbf{u}_{k\lambda} &= \mathbf{k} \times \mathbf{u}_{k\lambda}. \end{aligned} \tag{24–50}$$

Let us apply these relations to a pure, i.e., transverse, radiation field. We can choose a gauge in which $\phi = 0$, and thus take $\nabla \cdot \mathbf{A} = 0$, $\nabla \cdot \mathbf{E} = 0$. In that case $\mathbf{k} \cdot \boldsymbol{\varepsilon}_{\lambda k} = 0$, and hence the summation over possible polarizations will include only two components, $\boldsymbol{\varepsilon}_{k1}$ and $\boldsymbol{\varepsilon}_{k2}$, for each $\mathbf{k}$, where $\boldsymbol{\varepsilon}_{k\lambda}$ is perpendicular to $\mathbf{k}$. Thus

$$\nabla \times \mathbf{u}_{k1} = ik\mathbf{u}_{k2}. \tag{24–51}$$

For the first term in the Hamiltonian, we must evaluate

$$\int \pi^2\, dv = \frac{\epsilon_0}{L^3} \int \sum_k \sum_{k'} e^{i(\mathbf{k}+\mathbf{k}')\cdot\mathbf{r}} p_k p_{k'}\, dv$$

$$= \epsilon_0 \sum_k \sum_{k'} \delta(k + k') p_k p_{k'} = \sum_k \epsilon_0 p_k p_{-k}. \qquad (24\text{–}52)$$

Since $\boldsymbol{\pi}$ is real, Eq. (24–49) is equal to its complex conjugate, and since $\mathbf{u}_{k\lambda} = \mathbf{u}^*_{-k\lambda}$ we must have $p_k = p^*_{-k}$. Hence,

$$\frac{1}{\epsilon_0} \int \pi^2\, dv = \sum_k \sum_{\lambda=1}^{2} |p_{k\lambda}|^2. \qquad (24\text{–}53)$$

Also, from Eqs. (24–51) and (24–48),

$$\frac{1}{\mu_0} \int (\nabla \times \mathbf{A})^2\, dv = \sum_{k,\lambda} |ikq_{k\lambda}|^2 = \sum_{k,\lambda} k^2 |q_{k\lambda}|^2. \qquad (24\text{–}54)$$

Hence,

$$H = \tfrac{1}{2} \sum_{k,\lambda} (|p_{k\lambda}|^2 + k^2 |q_{k\lambda}|^2). \qquad (24\text{–}55)$$

Each term in Eq. (24–55) is the same as the Hamiltonian for a harmonic oscillator. Thus the equations of the electromagnetic field are equivalent to the equations of motion for a set of harmonic oscillators. The number of oscillators, $(\omega^2/\pi^2 c^3)\, d\omega$, corresponding to an angular frequency interval $d\omega$, results from taking the length of the box large in comparison with the wavelength of the radiation considered. By the application of statistical mechanics the mean energy of each oscillator when it is in equilibrium with walls at temperature T may be found. We have already seen that the resulting distribution law, Eq. (21–65), is in profound disagreement with experiment; it was Planck's effort to correct this treatment which originally led to the quantum hypothesis.

SUGGESTED REFERENCES

H. Goldstein, *Classical Mechanics*. Chapter 11 of this work forms an excellent introduction to the Hamiltonian formulation for fields. Our treatment of the transition from discrete to continuous systems parallels that in Goldstein's Section 11–1.

G. Wentzel, *Introduction to the Quantum Theory of Fields*. The *raison d'être* for the Lagrangian and Hamiltonian formulation of classical electrodynamics is to facilitate quantization, and discussions are rare except as a preliminary to quantum mechanical treatments. Wentzel's Chapter 1 is largely devoted to classical fields.

L. I. SCHIFF, *Quantum Mechanics*. The alternative treatments we have outlined are to be found in Chapters XIII and XIV of Schiff's very useful book on quantum mechanics.

EXERCISE

1. Show that the transformation of the Lagrangian density

$$\mathfrak{L}' = \mathfrak{L} + \frac{\partial}{\partial x^i}\left[C^i\left(\eta, \frac{\partial \eta}{\partial x^i}\right)\right],$$

where C^i is an arbitrary vector function of η and $\partial\eta/\partial x^i$, does not affect the Eulerian equation. Hence the choice of Lagrangian density to represent a given field is not unique.

APPENDIX I

UNITS AND DIMENSIONS IN ELECTROMAGNETIC THEORY

Classical mechanics is characterized by the fact that its mathematical formulation does not contain any fundamental constants inherent in the theory. Hence all physical laws in classical mechanics "scale" perfectly for any change in parameters over any arbitrary range of magnitudes. It is customary, although not at all mandatory, to formulate classical mechanics in terms of three-dimensional entities: mass (M), length (L), and time (T). The number could be increased, for instance, by choosing the constant K in the equation Volume $= K$ (Length)^3 to be different from unity and to have dimensions. The dimension L^3 is customarily identified with volume by choice, not by necessity. Similarly, the number of fundamental units can be decreased by arbitrarily defining certain constants to be unity and dimensionless. The convention $c = \hbar = 1$ frequently used in quantum mechanical calculations is such an example. In these units $L = T = M^{-1}$ arbitrarily.

We mention these examples only to indicate that the number of independent dimensions is arbitrary even in classical mechanics, although convenience suggests a specific choice. In general, the greater the number of dimensional entities chosen, the more independent units can be chosen to suit the orders of magnitude convenient for a particular purpose. It should be remembered, however, that changing units or even numbers of dimensions does not affect the physical content of any equation if it is correctly interpreted.

In classical mechanics the MLT system is used conventionally, and hence the issues discussed above are usually not of interest. For electromagnetic theory, the conventions are of more recent origin and appear more controversial.

Electromagnetic theory differs from classical mechanics by the fact that one constant, c, the velocity of light *in vacuo*, appears as a fundamental constant of the theory. Physical laws thus "scale" correctly over arbitrary magnitudes only if ratios of length and time are held constant. In this property electromagnetic theory exhibits a feature which special relativity extends to all laws of physics.

If the MLT system is used in the mechanical quantities in electromagnetic theory, the constant c having dimensions LT^{-1} will appear explicitly. Whether any additional dimensional units are introduced is entirely a matter of convention. As an example, if in Coulomb's law in the form $F = Kq_1q_2/r^2$ the constant K is chosen arbitrarily to be dimensionless, then the charge q automatically acquires the dimensions $L^{3/2}M^{1/2}T^{-1}$ and no basic units beyond those of M, L, and T need be specified. The justification for this procedure is analogous to that for setting K in the equation $V = KL^3$ equal to a dimensionless constant and thus giving a volume the dimension L^3.

If we do not choose a dimensionless constant in one of the equations relating mechanical and electromagnetic quantities, then we retain the freedom of choosing one of the electrical units arbitrarily and assigning to it a dimension. This has been done in what is called the mks system.

The particular set of units employed in this text is the mks system now in fairly general use. The principal convenience of this system is that it incorporates the common technical units—volt, ampere, coulomb, etc.—and is thus particularly suitable for treating applications that involve both "lumped" circuit parameters and fields. Since these technical units imply that the unit of time is the second and also define the power simply in watts, the natural choice of mechanical units are the meter, kilogram, and second. With this choice of electrical and mechanical units the constants ϵ_0 and μ_0 in the equations

$$\mathbf{F} = \frac{q_1 q_2}{4\pi\epsilon_0} \frac{\mathbf{r}_{12}}{r_{12}^3}, \tag{1}$$

$$\mathbf{F} = \frac{\mu_0}{4\pi} \iint \frac{\mathbf{j}_1 \times (\mathbf{j}_2 \times \mathbf{r}_{12})}{r_{12}^3}\, dv_1\, dv_2 \tag{2}$$

can be determined once we have selected a basic electrical unit. We shall postpone their numerical determination until we have seen how the electrical unit was chosen, but we note that two constants, ϵ_0 and μ_0, are carried in the equations when they are written in mks units, although only one constant, c, is fundamental to the theory.

Historically, a set of units (esu) was defined by using unity in place of $4\pi\epsilon_0$ in Coulomb's law, Eq. (1), and cgs units for mechanical quantities. This defines the electrostatic unit of charge with mechanical dimensions indicated above, and from this the units of potential, electric field, etc., are defined. On the other hand, if we set $\mu_0 = 4\pi$ in Eq. (2), and if cgs mechanical units are used, the equation defines a unit of current (with dimensions $M^{1/2}L^{1/2}T^{-1}$) called the electromagnetic unit (emu) or abampere. Units for other electrical quantities can be derived from the abampere and the cgs relations. The charge densities and current densities defined by Eqs. (1) and (2) respectively obey the relation

$$\mathbf{j}_{\text{emu}} = \rho_{\text{esu}} \frac{\mathbf{u}}{c}, \tag{3}$$

where c appears here as the measured ratio of the units. This ratio was first determined by Weber and Kohlrausch by measuring the discharge of a condenser whose electrostatic capacity was known. A consistent set of units (Gaussian units) is obtained by using the electrical quantities derived from ρ_{esu} and the magnetic quantities derived from $\mathbf{j}_{\text{emu}}$ and carrying c in the equations as the only constant.

As indicated in Eqs. (1) and (2), the mks system is commonly used in its rationalized form; the cgs units have been quoted in their unrationalized form. A rational system of units contains the factor 4π in Coulomb's and Ampere's laws of force so as to eliminate it in the Maxwell field equations which involve sources. General vector relations do not contain such factors. The appearance of the geometrical factor is indicated mathematically by the form of the Green's function. If a source (such as a point source) defines a problem in spherical symmetry, then in rational units one obtains 4π explicitly in the resulting solution; if the source (such as a line source) defines a field structure of circular symmetry, then 2π appears. A system of units analogous to the Gaussian system but in rationalized form is known as the Heaviside-Lorentz system.

Historically, the ampere was taken to be exactly $\frac{1}{10}$ of the abampere, or emu, of current. This fact enables us to determine the magnitude of μ_0 in Eq. (2), since we have seen that the abampere is defined by setting $\mu_0/4\pi = 1$. It is customary to use the coulomb, not the ampere, as the basic electrical quantity, and we need only transform the defining equation of the electromagnetic system with all quantities of unit size into mks units. Explicitly, in the electromagnetic cgs system,

$$1 \text{ dyne} = 1 \text{ (abampere)}^2$$

[since all lengths cancel on the right side of Eq. (2)], which in mks units for which 1 abampere = 10 coulomb/second, becomes

$$10^{-5} \text{ newton} \equiv 10^{-5} \frac{\text{kilogram-meter}}{\text{second}^2} = \frac{\mu_0}{4\pi} \frac{10^2 \text{ (coulomb)}^2}{\text{(second)}^2}$$

or

$$\mu_0 = 4\pi \times 10^{-7} \frac{\text{kilogram-meter}}{\text{(coulomb)}^2} = 4\pi \times 10^{-7} \frac{\text{henry}}{\text{meter}}.$$

The constant ϵ_0 is now obtained from the relation $\epsilon_0\mu_0 = 1/c^2$,

$$\epsilon_0 = \left(\frac{10^7}{4\pi c^2} \simeq \frac{1}{36\pi \times 10^9}\right) \left(\frac{\text{(coulomb)}^2 \text{ (second)}^2}{\text{kilogram (meter)}^3} = \frac{\text{farad}}{\text{meter}}\right).$$

What has been done here is to define μ_0 in terms of the arbitrarily chosen size and dimension of an electrical unit, whereupon ϵ_0 is automatically fixed if the system is to be consistent with the mechanical units and the experimental value of the fundamental constant c.

Electrical units which are referred to mechanical standards via defining relations containing fixed numerical constants are called *absolute* electrical units. Actually, the accuracy with which the emu and esu could be realized in terms of their defining equations was until recently insufficient for practical purposes: during the period when the accuracy of verification of absolute units was inferior to the reproducibility of standards, the practical units were based on such standards* as the *international* ampere and the *international* ohm. Improvements in techniques have resulted in greatly improved absolute electrical measurements, so that the former international standards have been relegated to the role of secondary standards. Hence the value $\mu_0 = 4\pi \times 10^{-7}$ as an exact numerical constant refers to the practical units as absolute rather than international units. Note that ϵ_0 in the mks system depends on the experimental relation between the velocity of electromagnetic radiation to the length and time standards (although μ_0 does not); this corresponds to the explicit presence of c in the Gaussian and Heaviside-Lorentz systems. In the "natural" system the velocity of light itself constitutes a standard. It is clear that an experimental measurement of the velocity of light can only provide a measure of the ratio of the velocity of propagation of electromagnetic radiation to the ratio of length and time standards. Hence the resultant number can never have any fundamental significance in an absolute sense, but is of great practical utility in providing independent accurate means of relating the length and time standards.

Clearly, the physical content of the fundamental relations is the same in all systems of units. It is easy to translate the laws of electrodynamics from one system to another: for vacuum conditions, the relations in this book are written so that the transformation

$$\begin{aligned} cB_{\text{mks}} &\to B_{\text{Gaussian}}, \\ \epsilon_0 &\to (4\pi)^{-1}, \\ \mu_0\epsilon_0 &\to 1/c^2 \end{aligned}$$

will effect a reduction to their Gaussian equivalents. Table I–1 contains a brief summary of fundamental electromagnetic relations valid *in vacuo* as they appear in the various systems of units. For convenience in the numerical conversion of units, a list of the most important conversion factors is given in Table I–2. Table I–4 contains other numerical constants and functional relations useful in applications involving atomic particles.

The situation regarding the equations in material media is somewhat more complex than that for vacuum conditions, and is frequently misunderstood. To gain some insight into the problem, let us consider the fundamental process by which Maxwell's equations in media are generated from the vacuum relations.

---

* For example, the international ampere was defined as "the value of the unvarying current which on passing through a solution of silver nitrate in water in accordance with standard specifications deposits silver at the rate of 0.001118 gm per second."

The vacuum source equations have the general form

$$D(F) = S, \tag{4}$$

where D is a linear differential operator acting on the field F and S is the source. In media S is broken up in terms of an "accessible" (macroscopic) source S and an "inaccessible" source S_p, i.e.,

$$D(F) = S + S_p. \tag{5}$$

S_p is then derived from an auxiliary quantity F_p by the same differential operator D such that

$$S_p = D(F_p) \tag{6}$$

and hence

$$D(F - F_p) = S. \tag{7}$$

A partial field $F_H = F - F_p$ can thus be defined such that

$$D(F_H) = S, \tag{8}$$

i.e., such that this field is derived from the "accessible" (often called "true") sources only. In Table I–3 this general statement is illustrated in terms of the actual electrodynamic quantities.

Ambiguity arises because in this general formulation quantities of the type F_p (i.e., **P** and **M**) appear in a dual role. On the one hand, in the relations

$$\rho_P = -\nabla \cdot \mathbf{P}, \tag{9}$$

$$\mathbf{j}_M = \nabla \times \mathbf{M}, \tag{10}$$

P and **M** represent purely source quantities—they describe certain charge and current distributions. In both relations, on the other hand, **P** and **M** can be viewed as fields, namely, those electric or magnetic fields whose sources are ρ_P or $\mathbf{j}_M$ respectively. In relations of the type

$$\epsilon_0 \mathbf{E} = \mathbf{D} - \mathbf{P}, \tag{11}$$

$$\mathbf{B}/\mu_0 = \mathbf{H} + \mathbf{M}, \tag{12}$$

as written in the now conventional mks system, **P** and **E** or **B** and **M** are given different units: **P** is measured in coulombs/meter^2, while **E** is in volts/meter; **M** is in amperes/meter, while **B** is in webers/meter^2. This convention emphasizes the roles of **M** and **P** as current and charge descriptions. The "partial field" aspect is, however, equally valid, and is somewhat obscured by the constants in Eqs. (11) and (12). In cgs units, Eqs. (11) and (12) become

$$\mathbf{E} = \mathbf{D} - 4\pi\mathbf{P}, \qquad \mathbf{B} = \mathbf{H} + 4\pi\mathbf{M} \text{ (Gaussian)},$$

$$\mathbf{E} = \mathbf{D} - \mathbf{P}, \qquad \mathbf{B} = \mathbf{H} + \mathbf{M} \text{ (Heaviside-Lorentz)}.$$

Here the units of all quantities are the same, and hence the "partial field' aspect is emphasized.

To solve field problems, it is in general necessary to specify the "constitutive equation" giving **P**(**E**) and **M**(**B**). If these are of linear form, such as

$$\mathbf{P} = \epsilon_0(\kappa - 1)\mathbf{E}, \tag{13}$$

it is often said that **E** represents an "intensive variable" and **P** an "extensive variable," i.e., **E** is cause and **P** effect. This point of view is emphasized by the fact that $\mathbf{E} \cdot \delta\mathbf{P}$ represents the differential of work done in this case. From this aspect the use of different units for **E** and **P** appears justified. Actually the cause-effect situation is very much less clear when permanent polarization (electrets) or permanent magnets are considered. It should be noted that the basic relations (11) and (12) are *additive;* on the other hand, relations of type (13) are not at all general.

We may summarize by saying that the question of whether **E**, **P**, and **D** (or **B**, **M**, and **H**) should have the same units is fairly irrelevant; in fact, an understanding of the *dual* physical function of **P** and **M** is the principal requirement for clarity in the classical theory of electric and magnetic media.

MISCELLANEOUS TABLES

Table I–1

FUNDAMENTAL ELECTROMAGNETIC RELATIONS VALID "IN VACUO" AS THEY APPEAR IN THE VARIOUS SYSTEMS OF UNITS

mks (rationalized)	Gaussian* (cgs)	Heaviside-Lorentz (cgs)	"Natural" units $c = \hbar = 1$ (rationalized)
$\nabla \cdot \mathbf{E} = \rho/\epsilon_0$	$\nabla \cdot \mathbf{E} = 4\pi\rho$	$\nabla \cdot \mathbf{E} = \rho$	$\nabla \cdot \mathbf{E} = \rho$
$\mathbf{E} = \frac{1}{4\pi\epsilon_0}\int \frac{\rho\mathbf{r}}{r^3}\,dv$	$\mathbf{E} = \int \frac{\rho\mathbf{r}}{r^3}\,dv$	$\mathbf{E} = \frac{1}{4\pi}\int \frac{\rho\mathbf{r}}{r^3}\,dv$	$\mathbf{E} = \frac{1}{4\pi}\int \frac{\rho\mathbf{r}}{r^3}\,dv$
$\nabla \times \mathbf{B} = \mu_0\left(\mathbf{j} + \epsilon_0\frac{\partial \mathbf{E}}{\partial t}\right)$ $\mathbf{B} = \frac{\mu_0}{4\pi}\int \frac{\left(\mathbf{j} + \epsilon_0\frac{\partial \mathbf{E}}{\partial t}\right) \times \mathbf{r}}{r^3}\,dv$	$\nabla \times \mathbf{B} = 4\pi\mathbf{j} + \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}$ $\mathbf{B} = \int \frac{\left(\mathbf{j} + \frac{1}{4\pi c}\frac{\partial \mathbf{E}}{\partial t}\right) \times \mathbf{r}}{r^3}\,dv$	$\nabla \times \mathbf{B} = \mathbf{j} + \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}$ $\mathbf{B} = \frac{1}{4\pi}\int \frac{\left(\mathbf{j} + \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}\right) \times \mathbf{r}}{r^3}\,dv$	$\nabla \times \mathbf{B} = \mathbf{j} + \frac{\partial \mathbf{E}}{\partial t}$ $\mathbf{B} = \frac{1}{4\pi}\int \frac{\left(\mathbf{j} + \frac{\partial \mathbf{E}}{\partial t}\right) \times \mathbf{r}}{r^3}\,dv$
$\nabla \cdot \mathbf{B} = 0$ $\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$	$\nabla \cdot \mathbf{B} = 0$ $c\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$	$\nabla \cdot \mathbf{B} = 0$ $c\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$	$\nabla \cdot \mathbf{B} = 0$ $\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$
$\mathbf{F} = e(\mathbf{E} + \mathbf{u} \times \mathbf{B})$	$\mathbf{F} = e\left(\mathbf{E} + \frac{\mathbf{u}}{c} \times \mathbf{B}\right)$	$\mathbf{F} = e\left(\mathbf{E} + \frac{\mathbf{u}}{c} \times \mathbf{B}\right)$	$\mathbf{F} = e(\mathbf{E} + \mathbf{u} \times \mathbf{B})$

$\mathbf{B} = \nabla \times \mathbf{A}$ $\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}$	$\mathbf{B} = \nabla \times \mathbf{A}$ $\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}$	$\mathbf{B} = \nabla \times \mathbf{A}$ $\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}$	$\mathbf{B} = \nabla \times \mathbf{A}$ $\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}$
$\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0$ $\nabla \cdot \mathbf{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t} = 0$	$\nabla \cdot \mathbf{j} + \frac{1}{c}\frac{\partial \rho}{\partial t} = 0$ $\nabla \cdot \mathbf{A} + \frac{1}{c}\frac{\partial \phi}{\partial t} = 0$	$\nabla \cdot \mathbf{j} + \frac{1}{c}\frac{\partial \rho}{\partial t} = 0$ $\nabla \cdot \mathbf{A} + \frac{1}{c}\frac{\partial \phi}{\partial t} = 0$	$\nabla \cdot \mathbf{j} + \frac{\partial \rho}{\partial t} = 0$ $\nabla \cdot \mathbf{A} + \frac{\partial \phi}{\partial t} = 0$
$F^{ij} = \begin{pmatrix} 0 & -cB_z & cB_y & +E_x \\ cB_z & 0 & -cB_x & +E_y \\ -cB_y & cB_x & 0 & +E_z \\ -E_x & -E_y & -E_z & 0 \end{pmatrix}$ $\frac{\partial F^{ij}}{\partial x^i} = \frac{j^j}{\epsilon_0}$	$F^{ij} = \begin{pmatrix} 0 & -B_z & B_y & +E_x \\ B_z & 0 & -B_x & +E_y \\ -B_y & B_x & 0 & +E_z \\ -E_x & -E_y & -E_z & 0 \end{pmatrix}$ $\frac{\partial F^{ij}}{\partial x^i} = 4\pi j^j$	F^{ij} has same form as in Gaussian units $\frac{\partial F^{ij}}{\partial x^i} = j^j$	F^{ij} has same form as in Gaussian units $\frac{\partial F^{ij}}{\partial x^i} = j^j$

* In some textbooks employing Gaussian units **j** is measured in esu. In that case the equations are those given here except that **j** appears with a factor $1/c$.

TABLE I–2

CONVERSION FACTORS

Multiply the number of mks units below	by	to obtain the number of Gaussian (cgs) units of
ampere	10^{-1}	current in abamperes
ampere/meter^2	10^{-5}	current density in abampere/cm^2
coulomb	3×10^9*	charge in esu
coulomb/meter^3	3×10^3*	charge density in esu/cm^3
farad = coulomb/volt	9×10^{-11}*	capacitance in cm
henry = volt sec/ampere	10^9	inductance in emu
joule	10^7	energy in ergs
newton	10^5	force in dynes
ohm = volt/ampere	$\frac{1}{30}$*	resistance in esu of potential per abampere
	$\frac{1}{9} \times 10^{11}$*	resistance in esu
volt	$\frac{1}{300}$*	potential in esu
volt/meter	$\frac{1}{3} \times 10^4$*	electric field intensity **E** in esu
coulomb/meter^2	$12\pi \times 10^5$*	electric displacement **D** in esu
weber = volt second	10^8	magnetic flux in maxwells
weber/meter^2	10^4	flux density **B** in gauss
ampere-turns/meter	$4\pi/10^3$	field intensity **H** in oersteds
mho/meter	$\frac{3}{10}$*	conductivity in abamperes/cm^2/esu of field intensity
ampere turns	$4\pi/10$	magnetomotive force in gilberts

* In all conversion factors marked by an asterisk 3 is used for $c/10^{10}$, where in this definition c is measured in cgs units. If higher accuracy is desired, a more precise value of c must be substituted. DuMond and Cohen give $c = 299792.9 \pm 0.8$ km/sec. [*Revs. Modern Phys.* **25**, 961 (1953).]

TABLE I-3

DEFINITION OF FIELDS FROM SOURCES (MKS SYSTEM)

	Electric	Magnetic	Equivalent covariant description
Vacuum (all sources "accessible")	$\nabla \cdot \mathbf{E} = \rho/\epsilon_0$	$\nabla \times \mathbf{B} = \mu_0\left(\mathbf{j} + \epsilon_0 \dfrac{\partial \mathbf{E}}{\partial t}\right)$	$\dfrac{\partial F^{ij}}{\partial x^i} = \dfrac{j^j}{\epsilon_0}$
Material media, sources separated	$\nabla \cdot \mathbf{E} = \dfrac{\rho + \rho_P}{\epsilon_0}$	$\nabla \times \mathbf{B} = \mu_0\left(\mathbf{j} + \mathbf{j}_M + \mathbf{j}_P + \epsilon_0 \dfrac{\partial \mathbf{E}}{\partial t}\right)$	$\dfrac{\partial F^{ij}}{\partial x^i} = \dfrac{1}{\epsilon_0}(j^j + j^j_M)$
Inaccessible sources defined from auxiliary function	$\nabla \cdot \mathbf{E} = \dfrac{\rho}{\epsilon_0} + \dfrac{(-\nabla \cdot \mathbf{P})}{\epsilon_0}$	$\nabla \times \mathbf{B} = \mu_0\left(\mathbf{j} + \nabla \times \mathbf{M} + \dfrac{\partial \mathbf{P}}{\partial t} + \epsilon_0 \dfrac{\partial \mathbf{E}}{\partial t}\right)$	$\dfrac{\partial F^{ij}}{\partial x^i} = \dfrac{1}{\epsilon_0}\left(j^j + \dfrac{\partial M^{ij}}{\partial x^i}\right)$
Definition of partial field	$\mathbf{D} = \epsilon_0 \mathbf{E} - (-\mathbf{P})$	$\mathbf{H} = \dfrac{\mathbf{B}}{\mu_0} - \mathbf{M}$	$H^{ij} = \epsilon_0 F^{ij} - M^{ij}$
Field equations in media	$\nabla \cdot \mathbf{D} = \rho$	$\nabla \times \mathbf{H} = \mathbf{j} + \dfrac{\partial \mathbf{D}}{\partial t}$	$\dfrac{\partial H^{ij}}{\partial x^i} = j^j$

TABLE I–4 USEFUL NUMERICAL RELATIONS

A. *Some atomic constants.**

cgs Gaussian	cgs Heaviside-Lorentz	mks	Name
$\dfrac{e^2}{\hbar c} = \dfrac{1}{137.04}$	$\dfrac{e^2}{4\pi\hbar c} = \dfrac{1}{137.04}$	$\dfrac{e^2}{4\pi\hbar c\epsilon_0} = \dfrac{1}{137.04}$	Fine structure constant
$\dfrac{e^2}{mc^2} = 2.8178 \times 10^{-13}$ cm	$\dfrac{e^2}{4\pi mc^2} = 2.8178 \times 10^{-13}$ cm	$\dfrac{e^2}{4\pi\epsilon_0 mc^2} = 2.8178 \times 10^{-15}$ m	Classical electron radius
$\dfrac{\hbar^2}{me^2} = 5.2917 \times 10^{-9}$ cm	$\dfrac{4\pi\hbar^2}{me^2} = 5.2917 \times 10^{-9}$ cm	$\dfrac{4\pi\epsilon_0\hbar^2}{me^2} = 5.2917 \times 10^{-11}$ m	Bohr radius

B. *Relations useful if energy of a particle is measured in electron volts.**

$E_0 = m_0c^2 = .51097$ Mev for electron

$E_0 = m_0c^2 = 938.210$ Mev for proton

1. "Magnetic rigidity" $B\rho$ of particle of kinetic energy T carrying charge e.

cgs Gaussian	*mks*
$B\rho = \dfrac{10^8}{c}\sqrt{T^2 + 2TE_0} \sim \dfrac{\sqrt{T^2 + 2TE_0}}{300}$	$B\rho = \dfrac{\sqrt{T^2 + 2TE_0}}{c} \sim \dfrac{\sqrt{T^2 + 2TE_0}}{3 \times 10^8}$

2. Tension τ of wire carrying current I having same orbit in magnetic field as particle of kinetic energy T carrying electronic charge e.

cgs Gaussian	*mks*
$\tau = I\sqrt{T^2 + 2TE_0} \times \dfrac{10^8}{c} \sim \dfrac{I\sqrt{T^2 + 2TE_0}}{300}$	$\tau c = I\sqrt{T^2 + 2TE_0}$

3. Velocity u, momentum p, kinetic energy T, total energy E.

$$\frac{E}{E_0} = \cosh\theta = \frac{T}{E_0} + 1 = \sqrt{\left(\frac{cp}{E_0}\right)^2 + 1} = \frac{1}{\sqrt{1 - (u^2/c^2)}}$$

$$\frac{cp}{E_0} = \sinh\theta = \sqrt{\left(\frac{T}{E_0}\right)^2 + 2\frac{T}{E_0}} = \sqrt{\left(\frac{E}{E_0}\right)^2 - 1} = \frac{(u/c)}{\sqrt{1 - (u^2/c^2)}}$$

$$\frac{u}{c} = \tanh\theta = \frac{\sqrt{2TE_0 + T^2}}{T + E_0} = \frac{\sqrt{E^2 - E_0^2}}{E} = \frac{cp}{\sqrt{(cp)^2 + E_0^2}}$$

$$\frac{T}{E_0} = \cosh\theta - 1 = \frac{1}{\sqrt{1 - (u^2/c^2)}} - 1 = \frac{E}{E_0} - 1 = \sqrt{\left(\frac{cp}{E_0}\right)^2 + 1} - 1$$

---

* The numerical values here are consistent with those given by J. W. M. DuMond and E. R. Cohen, *Revs. Modern Phys.*, **25**, 961 (1953).

APPENDIX II

VECTOR RELATIONS IN CURVILINEAR COORDINATES

A list of vector operators expressed in the common orthogonal curvilinear coordinates is often useful in the solution of physical problems. For the derivation of these relations, it is possible to procede quite formally from the definition of the operator ∇ in Cartesian coordinates and the transformation equations to other coordinate systems, but for physical applications it is advantageous to work from the geometrical definitions of gradient, divergence, and curl. One may first specify the coordinate system and derive the required expressions, or make a general derivation valid for any curvilinear coordinates and only then specify the coordinates. We shall follow the latter plan, first outlining a derivation valid for any right-handed system of orthogonal coordinates for which the line element is known, and then writing the particular forms for Cartesian, cylindrical, and spherical polar coordinates.

A line element in three dimensions is an infinitesimal displacement in space. If only one of three orthogonal coordinates q_1, q_2, q_3 is varied, the corresponding line element may be written

$$ds_1 = h_1\,dq_1, \tag{1}$$

together with similar expressions in q_2 and q_3. For any infinitesimal displacement,

$$ds^2 = h_1^2\,dq_1^2 + h_2^2\,dq_2^2 + h_3^2\,dq_3^2. \tag{2}$$

Now the gradient of a scalar function ψ is defined by the requirement that

$$\nabla\psi \cdot d\mathbf{s} = d\psi, \tag{3}$$

giving the change in ψ corresponding to the space displacement $d\mathbf{s}$. Then

$$(\nabla\psi)_1 = \lim_{ds_1 \to 0} \frac{\psi(q_1 + dq_1) - \psi(q_1)}{ds_1} = \frac{1}{h_1}\frac{\partial\psi}{\partial q_1} \tag{4}$$

is the general form of a gradient component.

To find the divergence, we shall consider an infinitesimal volume $dv = ds_1\;ds_2\;ds_3$ bounded by the surfaces $q_1 =$ constant, $q_1 + dq_1 =$ constant, etc., as indicated in Fig. II–1. Let us apply Gauss's divergence theorem, $\int \nabla \cdot \mathbf{A}\,dv = \int \mathbf{A} \cdot d\mathbf{S}$, to a vector $\mathbf{A}(q_1,q_2,q_3)$ with components A_1, A_2, A_3, integrating over this infinitesimal volume. The

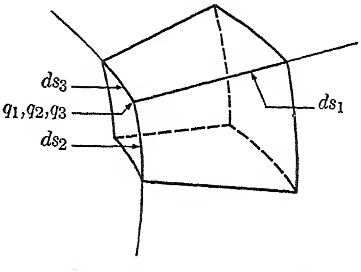


FIG. II–1 Element of volume for computing the divergence.

integral of the outward normal component of **A** over the two surfaces perpendicular to the direction of increasing q_1 is

$$\begin{aligned}(A_1\,ds_2\,ds_3)_{q_1+dq_1} - (A_1\,ds_2\,ds_3)_{q_1} &= \frac{\partial}{\partial q_1}(A_1\,ds_2\,ds_3)\,dq_1 \\ &= \frac{\partial(h_2h_3A_1)}{\partial q_1}\,dq_1\,dq_2\,dq_3 \\ &= \frac{1}{h_1h_2h_3}\,\frac{\partial(h_2h_3A_1)}{\partial q_1}\,dv,\end{aligned}$$

and analogous expressions hold for the other two sets of surfaces. Since the sum of these three terms is, by Gauss's theorem, equal to $\nabla\cdot\mathbf{A}\,dv$, the divergence is given explicitly by

$$\nabla\cdot\mathbf{A} = \frac{1}{h_1h_2h_3}\left(\frac{\partial(h_2h_3A_1)}{\partial q_1} + \frac{\partial(h_3h_1A_2)}{\partial q_2} + \frac{\partial(h_1h_2A_3)}{\partial q_3}\right). \tag{5}$$

The Laplacian of a scalar function can be written down immediately, since it is just the divergence of the gradient:

$$\nabla^2\psi = \frac{1}{h_1h_2h_3}\left[\frac{\partial}{\partial q_1}\left(\frac{h_2h_3}{h_1}\frac{\partial\psi}{\partial q_1}\right) + \frac{\partial}{\partial q_2}\left(\frac{h_3h_1}{h_2}\frac{\partial\psi}{\partial q_2}\right) + \frac{\partial}{\partial q_3}\left(\frac{h_1h_2}{h_3}\frac{\partial\psi}{\partial q_3}\right)\right]. \tag{6}$$

To obtain a particular component of the curl of a vector, we may apply Stokes' theorem to an infinitesimal area at right angles to the direction of the desired component. Consider the area defined by ds_2 and ds_3, as in Fig. II–2. By Stokes' theorem, $\oint \mathbf{A}\cdot d\mathbf{s} = \int(\nabla\times\mathbf{A})\cdot d\mathbf{S}$, which, in this application, becomes

$$(A_2\,ds_2)_{q_3} + (A_3\,ds_3)_{q_2+dq_2} - (A_2\,ds_2)_{q_3+dq_3} - (A_3\,ds_3)_{q_2} = (\nabla\times\mathbf{A})_1\,ds_2\,ds_3.$$

Therefore

$$(\nabla\times\mathbf{A})_1 = \frac{1}{h_2h_3}\left[\frac{\partial(h_3A_3)}{\partial q_2} - \frac{\partial(h_2A_2)}{\partial q_3}\right], \tag{7}$$

and the other two components are obtained by cyclic interchange of the coordinate indices.

In Table II–1 are listed the explicit forms for gradient, divergence, curl, and Laplacian in the three most common coordinate systems. For the definition of other orthogonal coordinates see, e.g., Margenau and Murphy, *The Mathematics of Physics and Chemistry*.

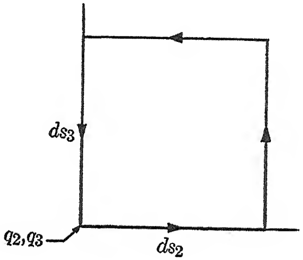


FIG. II–2 Element of area in the q_2, q_3 plane for finding the q_1 component of the curl. Arrows show the direction of the path of integration.

TABLE II–1

Cartesian coordinates	Cylindrical coordinates	Spherical polar coordinates
Orthogonal line elements		
dx, dy, dz	$dr, r\,d\theta, dz$	$dr, r\,d\theta, r\sin\theta\,d\varphi$
Components of gradient		
$(\nabla\psi)_x = \frac{\partial\psi}{\partial x}$ $(\nabla\psi)_y = \frac{\partial\psi}{\partial y}$ $(\nabla\psi)_z = \frac{\partial\psi}{\partial z}$	$(\nabla\psi)_r = \frac{\partial\psi}{\partial r}$ $(\nabla\psi)_\theta = \frac{1}{r}\frac{\partial\psi}{\partial\theta}$ $(\nabla\psi)_z = \frac{\partial\psi}{\partial z}$	$(\nabla\psi)_r = \frac{\partial\psi}{\partial r}$ $(\nabla\psi)_\theta = \frac{1}{r}\frac{\partial\psi}{\partial\theta}$ $(\nabla\psi)_\varphi = \frac{1}{r\sin\theta}\frac{\partial\psi}{\partial\varphi}$
The divergence of $\mathbf{A} = \nabla\cdot\mathbf{A}$		
$\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$	$\frac{1}{r}\frac{\partial(rA_r)}{\partial r} + \frac{1}{r}\frac{\partial A_\theta}{\partial\theta} + \frac{\partial A_z}{\partial z}$	$\frac{1}{r^2}\frac{\partial(r^2A_r)}{\partial r} + \frac{1}{r\sin\theta}\frac{\partial(\sin\theta A_\theta)}{\partial\theta} + \frac{1}{r\sin\theta}\frac{\partial A_\varphi}{\partial\varphi}$
Components of curl $\mathbf{A}$		
$(\nabla\times\mathbf{A})_x = \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right)$ $(\nabla\times\mathbf{A})_y = \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right)$ $(\nabla\times\mathbf{A})_z = \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right)$	$(\nabla\times\mathbf{A})_r = \left(\frac{1}{r}\frac{\partial A_z}{\partial\theta} - \frac{\partial A_\theta}{\partial z}\right)$ $(\nabla\times\mathbf{A})_\theta = \left(\frac{\partial A_r}{\partial z} - \frac{\partial A_z}{\partial r}\right)$ $(\nabla\times\mathbf{A})_z = \frac{1}{r}\left(\frac{\partial(rA_\theta)}{\partial r} - \frac{\partial A_r}{\partial\theta}\right)$	$(\nabla\times\mathbf{A})_r = \frac{1}{r\sin\theta}\left(\frac{\partial(\sin\theta A_\varphi)}{\partial\theta} - \frac{\partial A_\theta}{\partial\varphi}\right)$ $(\nabla\times\mathbf{A})_\theta = \frac{1}{r\sin\theta}\frac{\partial A_r}{\partial\varphi} - \frac{1}{r}\frac{\partial(rA_\varphi)}{\partial r}$ $(\nabla\times\mathbf{A})_\varphi = \frac{1}{r}\left(\frac{\partial(rA_\theta)}{\partial r} - \frac{\partial A_r}{\partial\theta}\right)$
Laplacian of $\psi = \text{div grad}\,\psi = \nabla^2\psi$		
$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}$	$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2\psi}{\partial\theta^2} + \frac{\partial^2\psi}{\partial z^2}$	$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\varphi^2}$

BIBLIOGRAPHY

WORKS DEALING PRIMARILY WITH CLASSICAL ELECTRODYNAMICS OR SOME OF ITS ASPECTS

Abraham, M., and R. Becker, *The Classical Theory of Electricity and Magnetism,* [from the 8th German ed. (1930)] Blackie and Son, 1932; [from the 14th German ed. (1949)] Hafner Publishing Co., 195–[?].

Becker, R., *Theorie der Elektrizität,* Band II, Teubner, 1933 (reprinted by J. W. Edwards, Ann Arbor, 1946).

Condon, E. U., Principles of Microwave Radio, *Revs. Modern Phys.* **14,** 341 (1942).

Coslett, V. E., *Bibliography of Electron Microscopy,* Arnold, 1950.

Harnwell, G. P., *Principles of Electricity and Electromagnetism,* 2nd ed., McGraw-Hill, 1949.

Hertz, H., *Electric Waves,* Macmillan (London), 1893.

Jeans, J. H., *The Mathematical Theory of Electricity and Magnetism,* 5th ed., Cambridge University Press, 1925.

Klemperer, O., *Electron Optics,* 2nd ed., Cambridge University Press, 1953.

Kraus, T. D., *Electromagnetics,* McGraw-Hill, 1953.

Lamont, H. R. L., *Wave Guides,* 3rd ed., Methuen, 1950.

Landau, L., and E. Lifshitz, *The Classical Theory of Fields* (trans. from the Russian), Addison-Wesley, 1951.

Lorentz, H. A., *The Theory of Electrons,* 2nd ed., Teubner, 1908 (reprinted by Dover, 1952).

Marton, C., and others, *Bibliography of Electron Microscopy,* U. S. Government Printing Office, 1950.

Mason, M., and W. Weaver, *The Electromagnetic Field,* University of Chicago Press, 1929.

Maxwell, J. C., *A Treatise on Electricity and Magnetism,* 3rd ed., Oxford University Press, 1904.

Planck, M. K. E. L., *Theory of Electricity and Magnetism,* Macmillan, 1932.

Rosenfeld, L., *Theory of Electrons,* North-Holland Publishing Co. (Amsterdam), 1951.

Schelkunoff, S. A., *Electromagnetic Waves,* Van Nostrand, 1943.

Schott, G. A., *Electromagnetic Radiation,* Cambridge University Press, 1912.

Slater, J. C., *Microwave Electronics,* Van Nostrand, 1950.

Slater, J. C., Microwave Electronics, *Revs. Modern Phys.* **18,** 441 (1946).

Slater, J. C., and N. H. Frank, *Electromagnetism,* McGraw-Hill, 1947.

Smythe, W. R., *Static and Dynamic Electricity,* 2nd ed., McGraw-Hill, 1950.

Sommerfeld, A., *Electrodynamics; Lectures on Theoretical Physics,* Vol. III, Academic Press, 1952.

Stratton, J. A., *Electromagnetic Theory,* McGraw-Hill, 1941.

Weber, E., *Electromagnetic Fields, Theory and Applications,* Vol. 1, *Mapping of Fields,* Wiley, 1950.

Zworykin, V. K., G. A. Morton, E. G. Ramberg, J. Hillier, and A. W. Vance, *Electron Optics and the Electron Microscope,* Wiley, 1945.

GRICULTURAL

MATHEMATICAL WORKS

Byerly, W. E., *Fourier's Series and Spherical Harmonics*, Ginn, 1893.

Churchill, R. V., *Fourier Series and Boundary Value Problems*, McGraw-Hill, 1941.

Courant, R., and D. Hilbert, *Methoden der Mathematischen Physik*, I, Springer, 1931; *Methods of Mathematical Physics*, Vol. I, Interscience Publishers, 1953.

Ince, E. L., *Ordinary Differential Equations*, Longmans, 1927 (reprinted by Dover, 1944).

Jahnke, E., and F. Emde, *Funktionentafeln mit Formeln und Kurven* (Tables of Functions), Teubner, 1933 (reprinted from 4th rev. ed. in English and German by Dover, 1945).

Kellogg, O. D., *Foundations of Potential Theory*, Springer, 1929 (reprinted by Frederick Ungar Publishing Co., n.d.).

MacRobert, T. M., *Spherical Harmonics*, 2nd ed., Dover, 1948.

Margenau, H., and G. M. Murphy, *The Mathematics of Physics and Chemistry*, Van Nostrand, 1943.

Morse, P., and H. Feshbach, *Methods of Mathematical Physics*, McGraw-Hill, 1953.

Pipes, L. A., *Applied Mathematics for Engineers and Physicists*, McGraw-Hill, 1946.

Synge, J. L., and A. Schild, *Tensor Calculus*, University of Toronto Press, 1949.

Watson, G. N., *Theory of Bessel Functions*, 2nd ed., Macmillan, 1944.

Weatherburn, C. E., *An Introduction to Riemannian Geometry and the Tensor Calculus*, Cambridge University Press, 1938.

Whittaker, E. T., and G. N. Watson, *A Course of Modern Analysis*, 4th ed., Cambridge University Press, 1927.

WORKS ON RELATIVITY THEORY

Bergmann, P. G., *An Introduction to the Theory of Relativity*, Prentice-Hall, 1942.

Einstein, A., *The Meaning of Relativity*, 4th ed., Princeton University Press, 1953.

Lorentz, H. A., A. Einstein, H. Minkowski, and H. Weyl, *The Principle of Relativity*, Methuen, 1923 (reprinted by Dover).

Møller, C., *The Theory of Relativity*, Oxford University Press, 1952.

Pauli, W., Relativitätstheorie, *Encyclopädie der Mathematischen Wissenschaften* V, Pt. 2, 539, Teubner, 1921.

Tolman, R. C., *Relativity, Thermodynamics and Cosmology*, Oxford University Press, 1934.

Tolman, R. C., *The Theory of the Relativity of Motion*, University of California Press, 1917.

WORKS CONTAINING MATERIAL ON OR RELATED TO CLASSICAL ELECTRODYNAMICS

Bergmann, P. G., *Basic Theories of Physics: Mechanics and Electrodynamics*, Prentice-Hall, 1949.

Blatt, J. M., and V. F. Weisskopf, *Theoretical Nuclear Physics*, Wiley, 1952.

Born, M., *Optik*, Springer, 1933.

Corben, H. C., and P. M. Stehle, *Classical Mechanics*, Wiley, 1950.

Fowler, R. H., *Statistical Mechanics*, 2nd ed., Cambridge University Press, 1936.

Goldstein, H., *Classical Mechanics*, Addison-Wesley, 1950.

Heitler, W., *The Quantum Theory of Radiation*, 3rd ed., Oxford University Press, 1954.

Joos, G., *Theoretical Physics*, 2nd rev. ed. (translated from the German ed. of 1934), Hafner Publishing Co., 1951.

Lamb, H., *Hydrodynamics*, 6th ed. rev., Macmillan, 1932 (reprinted by Dover, 1945).

Lindsay, R. B., and H. Margenau, *Foundations of Physics*, Wiley, 1936.

Lorentz, H. A., *Problems of Modern Physics*, Ginn, 1927.

Mayer, J. E., and M. G. Mayer, *Statistical Mechanics*, Wiley, 1940.

Morse, P. M., *Vibration and Sound*, 2nd ed., McGraw-Hill, 1948.

Page, L., *Introduction to Theoretical Physics*, 3rd ed., Van Nostrand, 1952.

Schiff, L. I., *Quantum Mechanics*, McGraw-Hill, 1949.

Slater, J. C., and N. H. Frank, *Introduction to Theoretical Physics*, McGraw-Hill, 1933.

Wentzel, G., *Quantentheorie der Wellenfelder* (Quantum Theory of Fields), F. Deuticke (Vienna), 1943 (reprinted by J. W. Edwards, 1946, and in English by Interscience Publishers, 1949).

NUTE • ALC

INDEX

Aberration, stellar, 237, 259, 348
Abraham, 350
Absorption of radiation, 333–336
Accelerated charge fields, 297–309
Accelerated frames of reference, 343
Accessible charge, 24, 379
Adiabatic theorem, 361
Advanced potential, 214
Ampère's law, 112, 376
Anisotropic media, 26
Associated Legendre functions, 74, 207
Astigmatic electron lenses, 361
Attenuation, wave guide, 198
Axial vector, 280

Bessel functions, 80, 138–140, 195, 308
  spherical, 202
Bessel's equation, 80, 189
Biot-Savart law, 114
Blackbody, 174, 209, 336
Boundary conditions, on electric field vectors, 26–28
  at conducting surfaces, 109
  on magnetic field vectors, 129
  on electromagnetic waves, 176–182
  at metallic boundaries, 185
Bragg, 332
bremsstrahlung, 303, 308
Brewster's angle, 179
Brillouin, 331

Canonical field variables, 369–373
Capacity, 56, 110
Cauchy-Riemann equations, 55
Cavities, oscillations in, 190–196, 202
Cavity, with "black" walls, 336
  definitions of electric field, 30
Čerenkov radiation, 309–312
Circulation density, 1
Classical electron theory difficulties, 320
Clausius-Mossotti relation, 34, 94, 101, 329
Collision, relativistic, 273
Complex fields, time average, 171
Complex transformations, 58–65, 136
Condensers, 110
Conductivity, 107
Conformal transformations, 58
Conjugate complex functions, 54–65
Conservation, of electromagnetic energy, 161
  of momentum, 162–164, 267
  in radiation reaction, 317
Conservation laws, 266, 270, 345, 349
Constitutive equations, 26, 129, 144, 350, 380
Continuity, hydrodynamic equations of, 93, 219
Contraction, Lorentz, 236, 248
  of a tensor, 264
Contraction hypothesis, 236
Contravariant tensor, 263
Convection potential, 293
Convective current, 118, 147, 342
Coulomb's law, 7, 86, 103, 376
Covariance, 261
Covariant electrodynamics, 339–350
Covariant mechanics, 261–273
Covariant tensor, 264
Cross section, for dipole absorption, 335
  for Rayleigh scattering, 327
  for scattering by a metallic cylinder, 201
    by a sphere, 207
  for Thomson scattering, 326
Curl, defined, 388
Current, convective, 118, 147, 342
  displacement, 121
  four-, 277, 339
  magnetization, 119
  polarization, 118, 176
  quasi-stationary, 152
  types of, 117
  vacuum displacement, 121
Current density, 107
Curvilinear coordinates, 287

Cutoff frequency, 196
Cylindrical cavities, oscillations in, 192–196
Cylindrical coordinates, 387–389
  Laplace's equation in, 80
  vector potential in, 138
  wave equation in, 195

D'Alembertian, 211, 277, 310
Damping, radiative, 322
Degeneracy of eigenvalues, 189
Delta function, 3, 22, 48, 84, 213, 310
Depolarization factor, 77
De Sitter, 329
Dielectric constant, 26
Dielectric cylinder and line charge, 66
Dielectric sphere and point charge, 75
Dielectric sphere in uniform field, 76
Diffraction, 332
Diffusion equation, 167
Dilation, time, 246
Dimensions, 376
Dipole, electric, interactions, 15
    radiation, 222
  magnetic, 193
    radiation, 226, 227
Dipole moment, 12
  volume distribution of, 20
Dipole potential, 13
Dirac delta function (*see* Delta function)
Dispersion, 328-331
Displacement current, 121
Displacement vector, 24
Divergence, defined, 387
  tensor, 96, 97, 265
Divergence theorem, 3
Doppler shift, 347
Dual of a tensor, 282

Eclipsing binary stars, 239
Eichenwald, 148
Eigenfunctions, 187
Einstein, 241, 261, 333, 351
Electric charge density, 8
Electric dipole, 15
  radiation, 222
Electric displacement vector, 25
Electric field, defined, 6
  in a medium, 29
Electric field lines, 23, 223
Electric flux, 55
Electric quadrupole, 13–15
Electric quadrupole radiation, 227, 228
Electric susceptibility, 25
Electric volume force, 92
Electromagnetic field tensor, 281, 341, 380
Electromagnetic mass, 314, 320, 349
Electromotive force, 108, 142, 155
Electron optics, 351-362
Electron radius, 88, 315, 325
Electret, 25, 380
Electrostatic lens, 358
Electrostatic potential, 9
Electrostriction, 95
"Emission" theories, 238
Energy density, in a dielectric, 89
  in electromagnetic field, 160
  in electrostatic field, 86
  in magnetic field, 153
  in radiation field, 174
Energy-momentum tensor, 345, 349
Energy velocity in a medium, 332
Equation of continuity, 93, 219
Equipotential surfaces, 55
Equivalence relations, 21
"Ether drag," 236
Ether frame, 231

Faraday disk, 344
Faraday's law of induction, 142
Ferromagnetic substances, 125
Feynman, 351
Field tensors, 341
Fizeau, 174, 176
Fluctuations, 333
Flux, of electric field, 55
  magnetic, 130, 142
Flux theorem, 7
Force, on current systems, 154
  in a dielectric liquid, 100
  on a dipole, 16
  Lorentz, 163, 172, 284, 345
  Minkowski, 272
  volume, 92, 159

Four-current, 277, 339
Four-momentum, 268, 347
Four-vector potential, 277
Fourier components, 212–228, 294, 304, 323, 331, 334
Free energy density, 90
Fresnel coefficient, 174–176, 259
Fresnel formulas, 178–179, 184
Functional derivative, 366

Galilean transformation, 230
Gauge transformation, 211, 278
Gaussian units, 301, 377
Gauss's divergence theorem, 3
  in tensor form, 96
Gauss's flux theorem, 7
Gedanken experimente, 245, 261, 273
Gradient, 265, 387
Green's function, for Laplace's equation 39, 45, 70, 84
  for the wave equation, 213
Green's reciprocity theorem, 38, 41
Green's theorem, 10, 50, 71
Group of Lorentz transformations, 253
Group velocity, 197, 331

Harmonic functions, 5, 65
Heaviside-Lorentz units, 377
Heitler, 317
Hertz, 225
Hertz potential, 219–228
Homogeneous wave equation, 166–209
Hughes, 129
Hydrodynamic equations of continuity, 93
Hysteresis, 154

Images, method of, 44, 69
Impedance, of free space, 168
  wave guide, 198
Index of refraction, 170, 179, 328, 330, 359
Inductance, 156
Induction, Faraday's law of, 142
Induction field, magnetic, 113
Induction zone, 225, 300
Infrared catastrophe, 305
Invariance, 252, 264
  gauge, 211, 278
  of phase, 347
Inversion, method of, 42

Joule heat losses, 155, 161

Kennedy, 236
Kennedy-Thorndike experiments, 240

Lagrangian, 352, 364–373
Langevin formula, 35
Laplace's equation, 9, 73, 133, 202
  in dielectrics, 38
Laplacian operator, 2, 133, 288, 389
Least action, principle of, 359
Legendre functions, 13, 74, 204, 207, 221
Lens, electron, 358–362
    astigmatic, 361
    electrostatic, 358
Liénard-Wiechert potentials, 280, 286–289, 292, 318
Light cone, 257
Linear multipole, 13, 221
Line element, special relativity, 252, 264
Lines, of charge, 12
  of force, 23
Lorentz condition, 210, 278
Lorentz contraction, 236, 248
Lorentz covariance, 261
Lorentz-Fitzgerald hypothesis, 236
Lorentz force, 163, 172, 284, 345
Lorentz frames, 343, 362
Lorentz transformation, 252, 253, 261, 283
Losses, cavity, 192
  guide, 198
  Joule heat, 155, 161
  at a metallic surface, 186

Magnetic circuits, 130
Magnetic dipole, 193
  radiation, 226
Magnetic energy density, 153
Magnetic field intensity, 124
Magnetic flux, 130, 142
Magnetic induction field, 113
Magnetic interaction of steady currents, 112
Magnetic moment, 119
Magnetic pole, 126

Magnetic scalar potential, 116
Magnetic susceptibility, 129
Magnetic volume force, 159
Magnetization, 119, 125, 150
Magnetization currents, 119
Magnetostatic lens, 359
Magnetostriction, 159
Majorana, 239
Mass, electromagnetic, 314, 320, 349
  variation with velocity, 268
Mass-energy relation, 269
Maxwell energy density, 87, 153
Maxwell equations, 143, 147, 166, 282
Maxwell stress tensor, 95, 162, 172, 345
Metallic boundary conditions, 185
Metric tensor, 264
Michelson and Morley, 176, 233
Miller, 234, 239
Minkowski diagram, 255, 340
Minkowski force, 272
Modes of oscillation, 192–196
Molar refraction, 329
Molecular fields, 31
Møller, 350
Momentum, electromagnetic, 162, 267, 315, 347
Motion of a conductor in a field, 148–151
Moving medium, Maxwell's equations in, 147
  plane waves in, 174
Multipoles, electrostatic, 12–15
  magnetic, 120
  radiation, 225–228
Mutual inductance, 156

Net point method, 52
Neumann's formula, 156
Nodal lines, in radiation field, 307

Octupole, 14
Oersted, 107
Ohm's law, 107, 108
Opalescence, 333, 337
Orthogonality of functions, 47, 50, 82, 189

Permanent magnets, 126
Permeability, 129
Permeable media, 129
Phase invariance, 347
Phase velocity, 170, 330
Planck's hypothesis, 305, 373
Plane of incidence, 176
Plane waves, 168–182
Plane wave expansion in spherical coordinates, 205
Poisson's equation, 9
Polar vector, 280
Polarizability, 33, 329
Polarization, volume, 20
Polarization charge density, 24
Polarization current, 118, 176
Polarization potential, 219
Poles, magnetic, 126
Postulates of special relativity, 241
Potential, advanced, 214
  convection, 293
  electrostatic, 9
  four-vector, 277
  Hertz, 219–228
  Liénard-Wiechert, 280, 286–289, 292, 318
  polarization, 219
  retarded, 214
  scalar, 4
    magnetic, 114, 131
  vector, 4
    of current distribution, 116
    magnetization, 120
Poynting vector, 161, 171, 198
Pressure, in dielectric liquids, 100–104
  radiation, 172
Propagation vector, defined, 170
Proper frame, 245
Proper time interval, 247
Pseudotensor, 280, 282

Q for cavity, 192
Quadrupole, oscillating, 227
Quadrupole lens, 361
Quadrupole moment, 13–15
Quasi-stationary currents, 152

Radiation, absorption of, 333–336
  from accelerated charge, 297–309
  Čerenkov, 309–312

Radiation (*continued*)
  dipole, 222, 226
  dipole antenna, 217
  dispersion of, 328–331
Radiation field, 300
Radiation pressure, 172
Radiation reaction, 314–320
Radiation resistance, 217
Radiation zone, 225, 300
Radiative damping, 322
Radius of electron, 88, 315, 325
Rasetti, 128
Rayleigh-Jeans law, 337
Rayleigh scattering, 327, 332
Reciprocity theorem, 38
Reflection, coefficient, 184
  metallic, 181
  of plane waves, 176
Refraction, 176
Refractive index, 170, 179, 359
Relativity, 150, 230–284, 339–350
Relaxation method, 52
Relaxation time, 111, 167
Reluctance, 130
Resistance, 108, 155, 217
Resonance, 324, 334
Retarded potential, 214
Ritz emission theory, 238
Roentgen, 148

Scalar potential, 4, 210
  electrostatic, 9
  magnetic, 114, 131
Scattering of radiation, by an electron, 325–330
  by a metallic cylinder, 199
  by a sphere, 203
Schiff, 345
Schwarz transformation, 59, 136
Self-inductance, 157
Separation of variables, 45–51, 65, 187, 201
Separation parameter, 46, 191, 195
Shankland, 234
Signal velocity, 331
Simultaneity, 243
Skin depth, 182, 187
Smolukowski, 333
Snell's law, 178
Sommerfeld, 331
Source density, 1
Space-time, average of electric field, 29
Space-time continuum, 254–257
Space-time interval, 252, 264
Specific inductive capacity, 26
Spherical Bessel functions, 202
Spherical cavities, 203
Spherical coordinates, 74, 201, 387–388
Spherical harmonics, 74
Spherical waves, 201
Stellar aberration, 237, 259, 348
Stewart, 239
Stokes' theorem, 9
Stratton, 133
Stream function, 55
Stress tensor, 95, 162, 172, 345
  at surface of dielectric, 98
"Strong focusing," 362
Substantial derivative, 94
Sum rule, 335
Summation convention, 15
Superposition principle, 1
Surface charge, 12, 18
Surface dipole layer, 17
Susceptibility, electric, 25
  magnetic, 129
Symmetry, of energy-momentum tensor, 349
Synchronization of clocks, 248

Telegraphers' equation, 170
Temperature dependence of dielectric constant, 34, 91
Tensor, contravariant, 263
  covariant, 264
  definitions, 263, 280
  dual of, 282
  electromagnetic field, 281, 380
  energy-momentum, 345, 349
  field, 341
  metric, 264
  stress, 95, 162, 172, 345
Tensor divergence, 96, 265
Tensor relations, 263–267
Tensor susceptibility, 26

Thomson moving-mirror experiment, 239
Thomson scattering, 326
Thomson's theorem, 91
Time average, complex vectors, 171
Time dilation, 247
Time of relaxation, 111, 167
Torque, on a dipole, 16
  on a moving dipole, 232
  on a moving lever, 272
Total reflection, 180
Trajectories of particles in fields, 355
Transformation relations, of current and charge, 277
  of electric and magnetic moment, 342
  of electromagnetic mass, 349
  of energy and momentum, 270
  of energy-momentum tensor, 348
  of fields, 283, 342
  of forces, 272
  of frequency and wave vector, 347
  of velocities, 257–258
Transformations, complex, 58
  Galilean, 230
  gauge, 211, 278
  Lorentz, 252, 261, 283
  multiple-angle, 64
  Schwarz, 59
  single-angle, 62
Transverse electric modes, 192
Transverse magnetic modes, 192
Transversality condition, 183
  on wave fields, 169
Trouton-Noble experiment, 232, 240, 294

Uniform field, potential of, 57
  potential of dielectric sphere in, 76
Unipolar induction, 240, 343
Uniqueness theorem for potentials, 5, 37, 132
Units, 375–386

Vacuum displacement current, 121
Variation principles, 355
Variational derivative, 366
Vector, covariant and contravariant, 262
  displacement, 24
  field, 1–6
  polar and axial, 280
Vector potential, 4, 126, 133, 210
Vector relations, 387–389
Virtual photons, 294
Volume force, electromagnetic, 345
  electrostatic, 92
  magnetic, 159

Wannier, 128
Wave equation, 166
Wave guides, 196
Wave guide attenuation, 198
Wave guide impedance, 198
Wave velocity, 170, 330
Weber and Kohlrausch, 376
Wheeler, 351
World line, 255, 351

X-ray yield, 308

Yukawa, 52, 369

ERRATA

for

CLASSICAL ELECTRICITY AND MAGNETISM

by PANOFSKY *and* PHILLIPS

(1*st edition*, 2*nd printing*)

ADDISON-WESLEY PUBLISHING COMPANY, INC.

READING, MASSACHUSETTS

Page 23, Problem 7: the upper limit of the second integral should be r''.

Page 44, line 7, *for* "conductor, and the potential and field on the left in" *read:* "conductor and the negative charge on the right removed; the potential on the right in Fig. 3–2(a)"

Page 56, line before (4–7), *for* "Eq. (4–5)" *read* "Eq. (4–6)."

Page 85, Problem 8, *for* "due to a disk" *read* "due to a conducting disk."

Page 113, Eq. (7–26), *for* "r^3" in the denominator, *read* "$r_{12}{}^3$."

Page 134, *delete from Eq.* (8–30) *through end of first paragraph,* p. 135, *and insert:*

$$\mathbf{A} = \nabla V + \nabla \times (\mathbf{a}U), \tag{8–30}$$

which obviously satisfies $\nabla \cdot \mathbf{A} = 0$. Here $\mathbf{a}$ is the vector of the (in general curvilinear) coordinate system which is to be chosen to fulfill certain conditions. The expression $\mathbf{B} = \nabla \times \mathbf{A}$ will not depend on V; hence U is the only function necessary to specify the field. This is as it should be, because in a source-free region there is no distinction between a field derivable from a scalar or vector potential. In general, however, V is needed to meet boundary value requirements on $\mathbf{A}$.

The function $\mathbf{A}$ as given by Eq. (8–30) satisfies the condition $\nabla \times (\nabla \times \mathbf{A}) = 0$ if $\mathbf{a}$ fulfills the condition that (in tensor notation) the function $\partial a_\beta/\partial x_\alpha$ is a diagonal matrix with equal coefficients, that is, $\partial a_\beta/\partial x_\alpha = \text{const}\ \delta_{\alpha\beta}$. Note that this condition implies that $\nabla \times \mathbf{a} = 0$. This condition is satisfied, for example, when $\mathbf{a} = \mathbf{x}_1$, $\mathbf{y}_1$, or $\mathbf{z}_1$, or by the vector $\mathbf{r}$ in spherical coordinates; the subscript-1 denotes a unit vector.

To prove this, we note that from (8–30),

$$\mathbf{B} = \nabla \times \mathbf{A} = -\nabla \times (\mathbf{a} \times \nabla U).$$

In tensor notation this becomes

$$-\nabla \times \mathbf{A}|_\alpha = \frac{\partial}{\partial x_\beta}\left[a_\alpha \frac{\partial U}{\partial x_\beta}\right] - \frac{\partial}{\partial x_\beta}\left[a_\beta \frac{\partial U}{\partial x_\alpha}\right],$$

GIFT

which can be reduced to

$$-\boldsymbol{\nabla} \times \mathbf{A}|_\alpha = -\frac{\partial}{\partial x_\alpha}\left[a_\beta \frac{\partial U}{\partial x_\beta}\right] + 2\frac{\partial U}{\partial x_\beta}\frac{\partial a_\beta}{\partial x_\alpha} - \frac{\partial U}{\partial x_\alpha}\frac{\partial a_\beta}{\partial x_\beta},$$

using the condition $\nabla^2 U = 0$, and the implied condition above. The first term of this relation is a pure gradient; the remaining terms are also pure gradients if $\mathbf{a}$ satisfies the condition above. Hence $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{A})$ vanishes.

Page 177, Eq. (11–41), numerator of second member, *for* "E" *read* "E''."

Page 180, Eq. (11–51), replace i/n^2 by i.

Page 180, Eq. (11–53), delete n^2 in the denominator.

Page 180, Eq. (11–54), replace i by i/n^2 in the first member. Read "$n^2 \cos\theta$" *for* "$\cos\theta$" in the denominator of the definition of ψ.

Page 182, second member of Eq. (11–61), for "E" *read* "E'."

Page 184, Problem 2, numerator of first equation, *for* "$2\cos\theta\cos\theta'$" *read* "$2\cos\theta\sin\theta'$."

Page 266, Fig. 16–1, *for* "d^4S_4 *read* "dS_4"; *for* "d^4S_4'" *read* "dS_4'."

Page 279, Eq. (17–16), *for* "$(x^j)_1 - (x^j)_2$" *read* "$(x^j)_2 - (x^j)_1$."

Page 308, Eq. (19–40), *for* "$e^{-i\omega t}$" *read* "$e^{+i\omega t}$."

Page 309, line below Eq. (19–43), *for* "per unit solid angle" *read* "into an angle $d\Omega$."

Page 309, line below (19–44), *for* "total rate of" *read* "total energy of the."

Page 319, Eq. (20–22), *for* "$\frac{r}{c}$" *read* "$\left(\frac{r}{c}\right)^2$."

AGRICULTURAL

Page 345, Eq. (22–20), *for* "$\partial T^{\circ}_{\alpha\beta}$" *read* "$\partial T^{M}_{\alpha\beta}$."

Page 356, left-hand side of Eq. (23–28), lower limit of integral, *for* "x_{α}" *read* "x_{α_1}."

Page 356, Eq. (23–30), limits on integral, for "τ_1" and "τ_2" *read* "x_1" and "x_2."

Page 370, equation below (17–30), right-hand member, *for* "$\mathbf{E} = \nabla \times \mathbf{A}$" *read* "$\mathbf{B} = \nabla \times \mathbf{A}$."

Page 372, Eqs. (24–50), *insert* "i" before "$\mathbf{k}$" in both.

Page 384, Table I–2 line 13 (volt/meter), should be $\frac{1}{3} \times 10^{-4}$.